MODELING COMMUNITY FIRST RESPONDER SYSTEMS: EVALUATING IMPACT, GUIDING RECRUITMENT, AND LEARNING DISPATCH STRATEGIES

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Hemeng Li

August 2024

© 2024 Hemeng Li
ALL RIGHTS RESERVED

MODELING COMMUNITY FIRST RESPONDER SYSTEMS: EVALUATING IMPACT, GUIDING RECRUITMENT, AND LEARNING DISPATCH STRATEGIES

Hemeng Li, Ph.D.

Cornell University 2024

In Community First Responder (CFR) systems, trained volunteers (CFRs) located near patients augment traditional emergency services by responding to alerts via a mobile app, especially for out-of-hospital cardiac arrests (OHCA) where rapid response is crucial. Volunteer efforts can significantly improve survival rates for OHCA.

It is important to determine the number of volunteers needed and the recruitment locations to achieve a target performance level. We first model CFR presence using a Poisson point process, allowing us to compute response-time distributions for the first-arriving CFR. Combining this model with known survival rate functions, we deduce survival probabilities for OHCA scenarios. Using convex optimization, we then determine the optimal distribution of CFRs across a region to optimize either the fraction of fast responses or the patient survival rate. This optimal CFR location distribution provides a benchmark for the best possible performance with a given number of volunteers, offering insights into the feasibility of introducing a CFR system in a new region or guiding additional recruitment in existing systems.

Additionally, we explore phased alerting policies for CFR systems, where volunteers are notified in stages with time delays, with the goal of maintaining high survival rates while minimizing so-called volunteer fatigue that can arise when

more than a required number of volunteers respond to a single OHCA. The policy defining these delays impacts both response times, directly related to survival, and the number of redundant (exceeding the required number) volunteer arrivals. We evaluate the performance of CFR dispatch policies through Monte Carlo simulation. We then present a Markov Decision Process (MDP) formulation and a machine learning-based selection strategy for determining which volunteers to alert and when for each incident, effectively balancing patient survival and volunteer fatigue.

We include a case study for both the CFR model and CFR dispatch strategies in Auckland, New Zealand, based on empirical data from their CFR system, GoodSAM. This comprehensive study provides a framework for improving CFR systems globally.

BIOGRAPHICAL SKETCH

Hemeng Li, originally from China, completed a Bachelor of Science degree with a joint major in Computer Science and Mathematics from Harvey Mudd College in 2019. Her research focuses on applied stochastic modeling, optimization, and data-driven decision making with applications in volunteer systems for out-of-hospital medical emergencies. She has provided practical insights on the evaluation of volunteer system impact, volunteer recruitment, and dispatch strategies. Her work has made significant contributions to the efficiency and effectiveness of emergency response systems. Through collaborations with organizations such as fire departments, the Red Cross, and ambulance providers in various countries, including New Zealand and the Netherlands, her research has had a tangible impact on improving public health and safety and will help save many lives.

This dissertation is dedicated to my parents and grandparents, whose love and support have been my constant source of strength and inspiration. Thank you for always being proud of me, for your unwavering belief in me, and for your endless encouragement throughout this journey.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Professor Shane G. Henderson, for his invaluable guidance, feedback, and support throughout my PhD journey. His dedication to research and teaching has always inspired me. He helped me identify and overcome my weaknesses, and without his help, this dissertation would not have been possible. I am also profoundly grateful to my committee members, Professor Jamol Pender and Professor Katya Scheinberg. You made my defense an enjoyable moment. Thank you for appreciating my work and for all your brilliant comments and questions.

I owe a special thanks to my wonderful collaborators, Professor Caroline Jagtenberg and Professor Pieter van den Berg. Working on this project with you was a pivotal experience in my PhD journey, and your support and collaboration made it possible. I particularly cherish the month we spent working together in person before my graduation. Thanks also to Océane Fourmentraux for her assistance in developing the volunteer dispatch with classification idea.

I would also like to acknowledge Bridget Dicker and Dave Richards from St John Ambulance Services in New Zealand for their helpful discussions and for providing crucial data for my research. This work was partially supported by the TKI Dinalog grant 2023-1-307TKI and the National Science Foundation grant CMMI-2035086. I am deeply appreciative of this financial support, which made my research possible.

To my friends Xumei Xi and Tonghua Tian, and all my PhD cohort at Cornell, your friendship will be treasured forever. The camaraderie and support, especially during those memorable first few semesters of classes, have been a highlight of my time here. The ORIE department has been like a second home to me. I am grateful to everyone in the department for making it a wonderful place to work

and study.

To my friend Peihan Zhao, thank you for being a wonderful friend since we were six, your support and friendship have been deeply appreciated. I appreciate that you were there to witness my PhD defense. And to my friend Wenjia Yan, thank you for always cheering me up and being there for me since high school. I truly appreciate you traveling to Ithaca for my graduation ceremony. Your presence made it much more memorable.

Finally, I would like to thank my family, who have constantly been my source of courage and motivation. I would not be where I am today without their belief in me and support, especially from my parents and grandparents. Your constant encouragement and faith in me have been my greatest strength.

To everyone who has been part of my journey, offering support, encouragement, and friendship along the way, this dissertation is a testament to all of you.

TABLE OF CONTENTS

Biographical Sketch . . . iii
Dedication . . . iv
Acknowledgements . . . v
Table of Contents . . . vii
List of Tables . . . x
List of Figures . . . xi

1 Introduction **1**

2 Modeling the Impact of Community First Responders **4**
2.1 Introduction . . . 4
2.2 Literature . . . 8
2.2.1 Performance Metrics . . . 8
2.2.2 Retrospective Studies on Effectiveness . . . 9
2.2.3 CFR for OHCA . . . 9
2.2.4 Dispatching . . . 10
2.2.5 Time-Sensitive Volunteers . . . 11
2.3 Modelling CFR Response . . . 12
2.3.1 Using a Poisson Point Process . . . 12
2.3.2 Response-Time Distribution . . . 13
2.4 The Ideal CFR Distribution . . . 17
2.4.1 Optimizing Coverage . . . 19
2.4.2 Optimizing Survival . . . 22
2.4.3 More Complex Survival Functions for OHCA . . . 23
2.4.4 Profile Optimization . . . 24
2.5 Case Description . . . 26
2.5.1 OHCA Incidence Rates . . . 27
2.5.2 Response Process . . . 27
2.5.3 CFR Profiles . . . 30
2.5.4 Performance Metrics . . . 31
2.6 Numerical Results . . . 32
2.6.1 CFR Density Requirements . . . 32
2.6.2 Quantifying the Benefit of a Given CFR Base . . . 34
2.6.3 Quantifying the Impact of CFRs' Geographical Spread . . . 36
2.6.4 Optimal Recruitment Efforts . . . 38
2.7 Conclusion and Discussion . . . 42

3 Phased Alerting of Community First Responders for Cardiac Arrest **46**
3.1 Introduction . . . 46
3.1.1 Background . . . 46
3.1.2 Importance . . . 47

3.1.3 Goals of This Investigation . . . 47
3.2 Methods . . . 48
3.2.1 Setting . . . 48
3.2.2 Study Design . . . 49
3.2.3 Policies . . . 49
3.2.4 Data . . . 50
3.2.5 Simulation . . . 52
3.2.6 Outcomes . . . 52
3.3 Results . . . 54
3.3.1 Parameter Estimation . . . 54
3.3.2 Performance of Policies . . . 55
3.4 Discussion . . . 56
3.4.1 Limitations . . . 58
3.4.2 Future Research . . . 58
3.4.3 Conclusions . . . 59

4 Learning to Dispatch Volunteers to Out-of-Hospital Medical Emergencies 64
4.1 Introduction . . . 64
4.2 Literature . . . 67
4.2.1 Community First Response for OHCA . . . 67
4.2.2 Volunteer Dispatch in Other Domains . . . 68
4.2.3 Machine Learning and Policy Selection for Real-time Decision Making . . . 70
4.3 Markov Decision Process Formulation . . . 71
4.3.1 State Space . . . 71
4.3.2 Action Space . . . 72
4.3.3 Transition . . . 72
4.3.4 Reward Function . . . 73
4.3.5 Optimal Dispatch Policy . . . 74
4.4 Selection Strategy Using a Classification Tree . . . 76
4.4.1 Dispatch Strategies and Pre-Defined Policies . . . 76
4.4.2 Simulation Model . . . 78
4.4.3 Machine Learning Pipeline . . . 79
4.5 Case Description . . . 80
4.5.1 CFR Response Process . . . 80
4.5.2 Empirical Data . . . 81
4.5.3 OHCA Incidents Generation . . . 82
4.5.4 Policies . . . 83
4.5.5 Performance Metrics . . . 83
4.6 Numerical Results . . . 84
4.6.1 High Interpretability of Tree-Based Strategy . . . 84
4.6.2 Adaptive Tree-Based Strategies vs. Static Strategies . . . 85
4.6.3 Impact of Volunteer Acceptance Rate . . . 89

4.6.4 MDP Optimal Policies vs. Tree-Selected Policies 91
4.7 Conclusion and Discussion 93

A Chapter 2 Appendix **97**
A.1 Convergence to a Poisson Point Process 97
A.2 Response Time Distributions 100
A.3 CFR Density Discretization 100
A.3.1 Data-Driven Experiment 100
A.3.2 Thought Experiment 103
A.4 Greedy Algorithm Pseudo-Code 104
A.5 Example with Two Locations 105
A.6 Frank-Wolfe Algorithm Pseudo-Code 106
A.7 Estimating the OHCA Incidence Rate 110
A.8 Retrospective Analysis 111
A.9 Profile Sensitivity Analysis 113
A.9.1 Impact of Profile 113
A.9.2 Uncertainty Analysis 113
A.9.3 Example with Larger Profile Impact 115
A.10 Homogeneous Ambulance Response 117
A.10.1 Quantifying the Benefit of a Given CFR Base 117
A.10.2 Quantifying the Impact of CFRs' Geographical Spread . . . 118
A.11 Ideal Recruitment 120
A.12 Profile Recruitment 121

B Chapter 4 Appendix **123**
B.1 Proof for MDP Optimal Policies 123
B.2 Numerical Results for Shorter EMS Response Times 129
B.2.1 Tree-Based Strategy with Shorter EMS Response Times . . 130
B.2.2 Adaptive Tree-Based Strategies vs. Static Strategies 131
B.2.3 Impact of High Volunteer Acceptance Rate 133

LIST OF TABLES

2.1 Required density of available CFRs (per km^2) to meet a range of response-time targets, assuming a 2-minute pre-trip delay and a walking speed of 6 km/h. 33

3.1 Results for $n = 10$ volunteers within 1 km, based on 1000 simulated volunteer locations, with 10000 view delay and acceptance simulations each. The maximum confidence interval halfwidth as a fraction of the mean among all these numbers was 0.012 (ignoring those cases where the mean was zero). Policies that send more than 10 alerts at time 0 are omitted, as there are only 10 eligible volunteers in this situation. 61

3.2 Results for $n = 30$ volunteers within 1 km, based on 1000 simulated volunteer locations, with 10000 view delay and acceptance simulations each. The maximum confidence interval halfwidth as a fraction of the mean among all these numbers was 0.009 (ignoring those cases where the mean was zero.) 62

3.3 Results for $n = 100$ volunteers within 1 km, based on 1000 simulated volunteer locations, with 10000 view delay and acceptance simulations each. The maximum confidence interval halfwidth as a fraction of the mean among all these numbers was 0.006 (ignoring those cases where the mean was zero.) 63

4.1 Accuracy and average objectives reported by the tree-based selection strategy versus the average objectives of the best static strategy among our predefined policies under a 14% volunteer acceptance rate. 86

4.2 Accuracy and average objectives reported by the tree-based selection strategy versus the average objectives of the best static strategy among our predefined policies under a 50% volunteer acceptance rate. 89

LIST OF FIGURES

2.1 Left: The empirical incident rate per area unit in Auckland, New Zealand. Right: Scatter plot of the actual and estimated number of cardiac arrests between 2013 and 2020 over all area units. . . . 28
2.2 Impact of increasing the CFR availability α on fraction of late arrivals and survival, assuming 1% of the population is signed up as CFRs, distributed over the area units proportionally to demand. . 35
2.3 Impact of increasing number of available CFRs ($n\alpha$) for different CFR distributions (ν) on the fraction of late arrivals ($\tau = 6$ min) and the survival probability. The black lines are each formed by the solution of 50 optimization problems. 37
2.4 Optimal CFR distribution (fraction of CFRs in that region divided by surface area) to minimize late arrivals for $\tau = 6$ min (a and c) or to maximize survival (b and d), with $n\alpha = 500$ or $n\alpha = 5000$. Darker colors represent higher CFR density. The white circles represent the location of the ambulance bases, and the numbers within them indicate the number of ambulances stationed at each base. . . 39
2.5 Time-dependent profile recruitment that maximizes survival for $n = 500/0.14 = 3571$. The white circles represent the location of the ambulance bases, and the numbers within them indicate the number of ambulances stationed at each base. 41

3.1 CFR response process. 48
3.2 Historical view delays and acceptance probabilities as observed on GoodSAM in New Zealand. 51
3.3 Distance-dependent travel times estimated based on historical data. 54

4.1 CFR response process. 65
4.2 Left: Histogram of triage delays from the 2825 OHCA incidents in Auckland from December 2017 to December 2020. Right: Histogram of view delays from the 6749 alerts in Auckland from December 2017 to December 2020 82
4.3 Tree-based selection strategy for instances in Auckland with 14% acceptance rate and 12-minute EMS response time measured from call arrival. 85
4.4 Trade-off between patient survival and number of redundant volunteers for static and adaptive tree-based strategies. Each blue dot represents the performance of one static strategy. The orange dot indicates the performance of the tree-based strategy. The green dot indicates the performance of the best static strategy, which is *Send 4 at time 0* under a 14% acceptance rate and a 12-minute EMS response time. 87

4.5 Percentage improvement in survival rate vs. the percentage improvement in the objective of the tree-selected policy compared to the best static policy for 2000 test cases under a 14% acceptance and a 12-minute EMS response time. 88
4.6 Trade-off between patient survival and number of redundant volunteers for static and adaptive tree-based strategies. Each blue dot represents the performance of one static strategy. The orange dot indicates the performance of the tree-based strategy. The green dot indicates the performance of the best static strategy, which is *Keep 2 active until 3* under a 50% acceptance rate and a 12-minute EMS response time. 90
4.7 Percentage improvement in survival rate vs. the percentage improvement in the objective of the tree-selected policy compared to the best static policy for 2000 test cases under a 50% acceptance and a 12-minute EMS response time. 92
4.8 Performance of MDP optimal policies vs. performance of tree-selected policies on 1000 test cases under a 12-minute EMS response time and a 14% acceptance rate within 15 minutes of CFR system activation. The blue dots represent instances where the tree-based strategy misclassifies the "best" policy among the pre-defined set. . 93

A.1 Cumulative distribution function (left) and density function of on-foot response time of closest CFR, depending on the expected number of available CFRs within 1 kilometer, for a response delay κ of 2 minutes, a walking pace w of 0.1 km/min and constant CFR density. These curves are also known as Rayleigh distributions with parameter $\sigma = 1/\sqrt{2\mu\pi}$. 101
A.2 Response time distribution under three different modeling assumptions. 102
A.3 Optimal distribution of CFRs over two points as a function of the fraction of demand coming from Location 1. Each curve corresponds to a single value of $\theta_l \in \{0.1, 0.5, 1, 2, 10, 100\}$ that is common to both locations. As we increase through these values, the curves approach the horizontal line at $\nu_1 = 0.5$. 107
A.4 Scatter plot of the actual number of cardiac arrests and the estimated number of cardiac arrests between 2013 and 2020 over all area units before (left) and after (right) adjusting the CBD and the Airport. 110
A.5 Histograms of distances from incidents to CFRs. Left: Distances of alerted CFRs from the 6749 alerts in Auckland since December 2017. Right: Distance of the *closest* alerted CFR for 194 incidents in the Auckland CBD since December 2017. The overlaid curve is the theoretical density function that follows from our model. 112

A.6 Impact of an increasing number of available CFRs for different CFR distributions (ν) on the fraction of late arrivals ($\tau = 6$ min) and the survival probability with EMS response for the optimal CFR distributions (ν) with and without profiles. Note that (a) and (b) have different scales. 114
A.7 Impact of inaccurate profile estimation for different recruitment distributions on the fraction of late arrivals ($\tau = 6$ min) and the survival probability with EMS response for the optimal CFR distributions (ν) with and without profiles. 115
A.8 Impact of inaccurate profile estimation for different recruitment distributions on the fraction of late arrivals ($\tau = 6$ min) and the survival probability with Fixed EMS response time for the optimal CFR distributions (ν) with and without profiles. 116
A.9 Impact of increasing the CFR availability α on fraction of late arrivals and survival, assuming 1% of the population is signed up as CFR, distributed proportionally over the area units. 117
A.10 Impact of increasing number of available CFRs ($n\alpha$) for different CFR distributions (ν) on the fraction of late arrivals ($\tau = 6$ min) and the survival probability. The black lines are each formed by the solution of 50 optimization problems. Note that (a) and (b) have different scales. 118
A.11 Optimal CFR distribution (fraction of CFRs in that region divided by surface area) to minimize late arrivals for $\tau = 6$ min (a and c) or to maximize survival (b and d), with $n\alpha = 500$ or $n\alpha = 5000$. Darker colors represent higher CFR density. 119
A.12 Optimal allocations of probability mass for late arrivals with different threshold times and survival, contrasted with a distribution of CFRs proportional to the demand density. Each mark represents one region (area unit): the horizontal axis shows the demand density per km^2 in that region; the vertical axis shows the fraction of CFRs in that region divided by its area. 120
A.13 Scatter plots of (a) the optimal number of available CFRs recruited and (b) the resulting CFR densities (available CFRs per square km), with and without profiles, when maximizing survival for 500 available CFRs. Each point represents one area unit and in both figures, the CBD is highlighted in red. 122

B.1 Tree-based selection strategy for instances in Auckland with 14% acceptance rate and 6-minute EMS response time measured from call arrival. 130

B.2 Trade-off between patient survival and number of redundant volunteers for static and adaptive tree-based strategies under a 14% acceptance rate. Each blue dot represents the performance of one static strategy. The orange dot indicates the performance of the tree-based strategy. The green dot indicates the performance of the best static strategy, which is *Send 1 at time 0* under a 6-minute EMS response time (left) and *Send 3 at time 0* under a 9-minute EMS response time (right). 132
B.3 Percentage improvement in survival rate vs. the percentage improvement in the objective of the tree-selected policy compared to the best static policy for 2000 test cases under a 14% acceptance and a 6-minute EMS response time. 133
B.4 Trade-off between patient survival and number of redundant volunteers for static and adaptive tree-based strategies under a 50% acceptance rate. Each blue dot represents the performance of one static strategy. The orange dot indicates the performance of the selected tree-based strategy. The green dot indicates the performance of the best static strategy, which is *Send 2 at time 0* under a 6-minute EMS response time (right) and *Send 3 at time 0* under a 9-minute EMS response time (left). 134
B.5 Percentage improvement in survival rate vs. the percentage improvement in the objective of the tree-selected policy compared to the best static policy for 2000 test cases under a 50% acceptance and 6-minute EMS response time. 135

CHAPTER 1

INTRODUCTION

Supplementing traditional ambulance services with community first responders (CFRs) can be highly beneficial for medical emergencies that require a fast response. CFR systems are particularly useful for out-of-hospital cardiac arrest (OHCA), a leading cause of mortality worldwide. Early access to treatment, such as cardiopulmonary resuscitation (CPR), significantly increases OHCA patient survival (Berdowski et al. 2010, Nichol et al. 1999). Applications like GoodSAM Platform (2020) and PulsePoint (2020) have been implemented in multiple countries, including the UK, New Zealand, and the USA, to facilitate CFR deployment for OHCA. Although there are CFR systems that address general medical emergencies, this dissertation focuses specifically on OHCA applications.

CFR systems work together with traditional emergency medical services (EMS). When an emergency occurs, the call center identifies registered CFRs near the incident. In early CFR systems, CFRs were dispatched based on their home and work addresses and were alerted through text messages (HartslagNu 2020), but with the widespread use of smartphones, CFR locations can now be tracked and alerted via GPS-enabled applications. This technology allows dispatchers to alert selected CFRs in real time. CFRs who choose to respond provide first aid until the ambulance arrives.

Our work examines several critical aspects of CFR systems. One key focus is understanding how CFR availability and their geographic distribution impact response times and patient outcomes. Traditional methods of analyzing CFR systems (Ringh et al. 2011, Zijlstra et al. 2014, Stieglis et al. 2020, Pijls et al. 2016, 2019) are restricted to data analyses, and therefore are limited by data sparsity, because

OHCA incidents are rare and thus data sets are small, even after years of collecting data. Therefore, we propose a mathematical modeling approach using a spatial Poisson point process model to estimate CFR response time distributions based on CFR densities. This model can be applied to both existing and prospective CFR systems and, combined with convex optimization, provides valuable insights into optimal CFR recruitment strategies.

Another important consideration is managing the inherent uncertainty in CFR alert acceptance. Ideally, notifying all nearby CFRs would minimize response times and improve patient survival, but this can lead to volunteer fatigue from excessive alerts or from multiple volunteers responding to a single OHCA, reducing future participation rates. Some CFR systems use phased alerting policies, sending alerts in batches with time delays, but these policies do not always account for real-time CFR locations or volunteer fatigue. To address these issues, we introduce both a Markov Decision Process (MDP) formulation and a machine learning-based selection strategy. These methods determine which volunteer to alert and when for each incident, balancing patient survival with volunteer fatigue. Our approach also leverages real-time information on CFR locations and dispatched ambulances unlike current practices which uses a fixed policy for all incidents.

The contributions of this dissertation are two-fold: introducing a new class of CFR problems to the operations research community, and developing and demonstrating methods that can inform practical decision-making to save lives. The dissertation is structured as follows:

- Chapter 2: Joint work with Pieter van den Berg, Shane Henderson, and Caroline Jagtenberg, discussing the use of a spatial Poisson point process model and convex optimization to evaluate CFR impact and guide recruit-

ment. This chapter has been accepted for publication at *Management Science* (van den Berg et al. 2024a).

- Chapter 3: Joint work with Pieter van den Berg, Shane Henderson, Caroline Jagtenberg, and Bridget Dicker, using Monte Carlo simulation to evaluate the impact of CFR dispatching policies on patient outcomes. This chapter has been submitted for publication (van den Berg et al. 2024b).
- Chapter 4: Joint work with Pieter van den Berg, Océane Fourmentraux, Shane Henderson, and Caroline Jagtenberg, developing an MDP formulation and a tree-based selection strategy to optimize CFR dispatch.

CHAPTER 2

MODELING THE IMPACT OF COMMUNITY FIRST RESPONDERS

2.1 Introduction

Certain medical emergencies require such a fast response that it can be helpful to supplement traditional ambulance services with community first responders (CFRs). CFRs are volunteers available to be dispatched by an ambulance control center, and can be members of the public who have received training or off-duty medical professionals.

The predominant application is out-of-hospital cardiac arrest (OHCA), for which in Europe alone 19 of 29 countries have a CFR system in place (Oving et al. 2019). OHCA is a major cause of mortality around the world, and the probability of survival significantly improves if patients have early access to treatment such as cardiopulmonary resuscitation CPR (Berdowski et al. 2010, Nichol et al. 1999). Alternative applications where CFRs are dispatched to general medical emergencies are surveyed in Phung et al. (2017) and include Lives (n.d.), NHS North West Ambulance Service (n.d.) and Northern Ireland Ambulance Service Health & Social Care Trust (n.d.).

CFR systems are activated by an ambulance control center in parallel with traditional emergency medical services (EMS). In the past, CFRs were dispatched based on their self-reported home and work address (HartslagNu 2020), while, in recent years, an increase in smartphone usage has enabled applications (apps) that alert CFRs based on their real-time GPS location. Such systems include PulsePoint

in the USA (PulsePoint 2020) and GoodSAM in several countries including the UK, Australia, and New Zealand (GoodSAM Platform 2020, Smith et al. 2017). These systems monitor the real-time location of CFRs who have the app running on their smartphones, though this data is not logged for privacy reasons. CFRs near an incident are alerted and some random subset of them accept the alert and proceed to the scene. CFRs typically do not have trouble accessing the patient, as the person who made the emergency phone call is informed about their dispatch. The CFR provides first aid to the level that their training allows, until the ambulance crew, which is dispatched independently, takes over.

The central goal of this chapter is to determine how the availability of CFRs and their distribution throughout the city affect response times. Faster responses are often equated with better care and consequently, response times are commonly used as a proxy for a system's effectiveness. The response time largely depends on the CFR density, which in turn depends on the number of registered CFRs and their distribution over the city. The number of registered CFRs ranges from 0.1% of the population in a newly introduced system in New Zealand (Global Resuscitation Alliance 2019) to more than 1% in a mature system in the Netherlands (Wikipedia 2020). Another major factor is the acceptance probability of CFRs; Brooks et al. (2016) found that PulsePoint CFRs accepted on average 23% of the alerts they receive, though this value varies considerably from one city to another. One might be concerned that CFRs could receive simultaneous alerts, leading to queueing effects, but the rate at which CFRs receive alerts is typically very low, e.g., Pijls et al. (2019) reports an average of 1.3 alerts per CFR per year.

One might attempt to understand CFR response by looking at how recorded CFR response times differ as a function of location. Doing so is unlikely to give

good results because of data sparsity. Indeed, the number of incidents per location is very small, leading to few observations even after collecting years of data. Even if such data were sufficient to accurately estimate how the number of incidents differs from location to location, much more data would be needed for predicting the distribution of response times.

In contrast with these retrospective approaches, we offer a prospective method, by estimating the performance of CFR systems using a stochastic model. Besides being less sensitive to statistical error, an additional benefit of this method is that it can be applied to either existing or prospective CFR systems. The model we propose uses a Poisson point process to estimate response time distributions based on CFR densities and is combined with traditional EMS response and survival functions from literature. The Poisson point process model enables the use of convex optimization to obtain bounds on the potential performance of various CFR deployments. We can then answer the following questions:

1. In what areas is introducing a CFR app an efficient way to reach certain response-time targets? (Section 2.6.1)
2. What is the benefit of introducing a CFR app for a base of already trained volunteers, with known home addresses? (Section 2.6.2)
3. What is the benefit of a CFR app with a given number of registered volunteers, when you do not know where they live or spend their time? (Section 2.6.3)
4. Where should one recruit additional CFRs for an existing app with known current CFR base? (Section 2.6.4)

All of these questions arose in real-life use cases. Question 1 was of interest

to the Fire Department of Amsterdam, which adopted a moonshot goal of having a CFR app and a large CFR base by 2030 (Brandweer Amsterdam-Amstelland 2021, p. 6). Question 2 was posed by a national Red Cross organization that considered extending the use of its existing volunteer base via an app. Both of these parties executed a case study based on the models in a preliminary version of this chapter. Questions 3 and 4 are relevant to the St. John Ambulance Service that operates GoodSAM in Auckland and we performed corresponding calculations that are partially included in this chapter.

We view the primary contributions of this work to be 1) the introduction of a new class of CFR problems to the Operations Research community, and 2) the selection of, and interplay between, Poisson point processes to model volunteers and convex optimization to perform spatial optimization. The combination of these techniques enables powerful and transparent analyses that can inform, and have informed, decisions that save lives.

The remainder of this chapter is structured as follows. Section 2.2 reviews the existing literature and Section 2.3 introduces our model that quantifies the impact of a *given* CFR base. Building upon this model, Section 2.4 introduces optimization models to determine the ideal geographical distribution of CFRs. We demonstrate our methods within the city of Auckland, New Zealand in Sections 2.5 and 2.6. We conclude by answering the four questions above and discussing the managerial implications of this work in Section 2.7. Appendix A provides supporting details.

2.2 Literature

We discuss performance metrics for CFR systems in Section 2.2.1 and review studies that evaluate the effectiveness of such systems in Section 2.2.2. Section 2.2.3 summarizes CFR literature that is specific for OHCA patients and Section 2.2.4 reviews studies on first-responder dispatching. Section 2.2.5 places our work in a broader context.

2.2.1 Performance Metrics

As an essential goal of CFR systems is to reach patients quickly, their performance can be measured in terms of response times: the duration between the moment a call arrives in the emergency call center and the moment the first responder arrives on scene. The same is true for EMS providers, who often have targets that are a function of response times, phrased as service level agreements (respond to $x\%$ of all calls in y minutes).

One way of translating response times into a medical outcome-based metric is to convert them to a survival probability through a so-called survival function. This is highly dependent on the medical condition of the patient. Survival functions have been used in work on ambulance operations in the absence of volunteer schemes; see Erkut et al. (2008), McLay (2009), Bandara et al. (2012), Zaffar et al. (2016).

2.2.2 Retrospective Studies on Effectiveness

In contrast to our work, existing literature focuses on retrospective analysis of data to determine the impact of CFR systems.

Early work from Sweden (Ringh et al. 2011) analyzed the impact of a CFR app using both a physical experiment and a small retrospective study in which volunteers arrived before ambulances approximately half the time. For a Dutch text-based volunteer CFR system, Zijlstra et al. (2014) compared response times in a system with volunteers to a system with ambulances only, reporting a reduction of 159 seconds in the time to defibrillation. Pijls et al. (2016) focused on survival rates: the survival rate for OHCA increased from 16% to 27% when at least one volunteer responded.

Several studies specifically investigated the impact of volunteer density. Both Stieglis et al. (2020) and Jansma (2014) investigated how density relates to response times via an empirical study and computer simulation, respectively. Pijls et al. (2019) show a positive correlation between the fraction of inhabitants registered as volunteers and patient survival.

2.2.3 CFR for OHCA

Some CFR systems are designed explicitly for OHCA patients, perhaps because OHCA survival is known to significantly improve with early CPR administered by trained individuals (Nichol et al. 1999, Sasson et al. 2010, Yan et al. 2020). This relationship between time to CPR and survival has been made explicit in multiple studies which have followed a group of patients for which the time to CPR has

been recorded, e.g., Valenzuela et al. (1997), Waalewijn et al. (2001), and De Maio et al. (2003).

Cardiac arrest patients are known to benefit from receiving defibrillation by automated external defibrillators (AEDs). CFR systems are therefore often designed to include information on AED locations. The optimization of AED placement is explored in Folke et al. (2009), Chan et al. (2016, 2018); however, they treat this question without explicitly incorporating the dynamics of a CFR system. Like those studies, we view as out of scope the detailed modeling of AED-CFR dispatch coordination, such as when one CFR is sent directly to the patient while another is sent to retrieve the nearest AED. While such considerations are important for an individual OHCA, such detailed modeling seems unnecessary for the high-level questions of recruitment that we consider here.

2.2.4 Dispatching

At the time of an incident, real-time decisions must be made on which volunteers are dispatched by sending a push notification to their phone. This decision is typically based on the observed locations of nearby volunteers and should balance overburdening volunteers with response-time benefits (Henderson et al. 2022).

Advanced dispatch methods may vary the number and time of alerts, and may consider a range of tasks, e.g., picking up medical equipment (Nazarian 2018, Matinrad et al. 2019). In this context, Matinrad et al. (2021a) discusses uncertain task compliance.

A recent trend is to have medical supplies delivered by drones, for example AEDs (Boutilier et al. 2017, Boutilier and Chan 2022, Chu et al. 2021) or

blood (Nisingizwe et al. 2022). Such a design disconnects equipment delivery decisions from CFR coordination, hence further confirming our choice to leave such questions out of scope for this chapter.

Recent unpublished work (Liu et al. 2022, Shin et al. 2022) considers AED drone management under ambulance as well as bystander response. Both papers model bystander availability through a Bernoulli distribution and the bystanders' response time does not depend on the number of available responders. This work, besides evincing very recent interest in the topic, highlights the unavailability of a detailed CFR response model, a gap that we aim to fill here.

2.2.5 Time-Sensitive Volunteers

More broadly, our work relates to the literature on time-sensitive volunteers and crowdsourcing, e.g., Manshadi and Rodilitz (2022), Ata et al. (2019a), McElfresh et al. (2020). The key differences with our work are that, in that setting, 1) requests that are relevant for any fixed volunteer are likely to be more frequent; 2) volunteer disengagement is a central consideration and is modeled in a variety of ways; and 3) the proximity of volunteers is relevant, but not critical as in our setting. There is a slight resemblance between our work and the literature on matching and crowdsourcing, e.g., Johari et al. (2021), Tafreshian et al. (2020), Özkan and Ward (2020); queueing, pricing and learning are central in this literature but not in our work. The use of spatial Poisson processes to model the location of servers (CFRs in our setting) could potentially be relevant in many applications, including the modeling of available drivers in ride hailing; see, e.g., Qin et al. (2020).

2.3 Modelling CFR Response

In this section, we argue that CFR locations are well modeled by a Poisson point process and thereby derive the distribution of their response times. This distribution, particularly when complemented with the distribution of EMS response times, can be translated to an expected health gain in the population due to CFRs.

Throughout the chapter we use the term "city" to refer to the overall area in which the CFR response is modeled. However, our modeling approach applies whether the area under consideration is a city, a county, a state, or even a nation, at least in principle.

2.3.1 Using a Poisson Point Process

We say that a CFR is *available* if they are present in the city, have the app running on their phone, and will accept a notification should it be sent to them. We assume that available CFRs are distributed throughout the city according to a spatial Poisson point process. We assume that the set of available CFRs does not depend on call volume, which is reasonable since typically CFRs are called out on the order of once per year (Pijls et al. 2019), at least in the context of OHCA.

Modeling available CFRs as a spatial Poisson point process is reasonable because of results that justify approximating certain spatial point processes by Poisson point processes. For a general introduction to Poisson point processes, including theoretical results that justify Poisson point process modeling in applications in great generality, see Kingman (1993) and Barbour et al. (1992). We justify the Poisson point process assumption with Proposition A.1.1 in Appendix A.1, which

shows that a Poisson point process arises in a regime where the number of CFRs n is large. That result is not the most general result possible, nor will it be surprising to those versed in point-process theory, but we provide it to support our contention that the Poisson point process model is a good one in our setting. This assertion is further supported by the prevalence of Poisson point process modeling in related settings; see Larson and Odoni (1981), for example.

Proposition A.1.1 shows that even when CFRs have a unique availability probability and location distribution, the location of available CFRs is well modeled by a Poisson point process with a certain mean measure μ. Thus it is not necessary to assume CFRs are a homogeneous group: each responder may have their own probability of accepting alerts and their own distribution of time spent in each part of the city. The exact way in which individual CFRs contribute to the aggregate mean measure μ is not important for our analysis; the overall measure μ is what is important.

It is therefore sufficient to henceforth assume that $\mu = n\alpha\nu$, where n indicates the overall number of CFRs, who have the *same* availability probability $\alpha \in (0, 1)$ and the *same* location probability distribution ν conditional on being available.

2.3.2 Response-Time Distribution

The response time to an incident is the minimum of the time till a CFR or ambulance arrives at the scene, with these intervals measured from the moment the call to EMS is initiated.

As is common in CFR apps (Smith et al. 2017), we assume that the closest available CFR is dispatched to an incident, as long as the CFR is within a given

maximal distance from the patient. We are interested in the distribution of the response time of this closest CFR, which depends on two components: 1) the response delay, and 2) walking time. The response delay is the time that passes between the call initiation and the responder starting to travel. This consists of the interval between the time the call is made and the time a CFR is dispatched (triage and dispatch delay), and the interval between the CFR dispatch and the time a CFR starts walking (CFR acceptance delay). For simplicity, the total response delay for a CFR is assumed to be constant and is denoted by κ. For any location x, define the ball $B(x,t)$ as the area surrounding x within which a CFR can reach the location within t minutes, including response delay.

We assume that ambulance bases are given, and ambulances respond from their bases and not from the road. This is an approximation that is reasonable when the ambulances are not too heavily loaded. Thus, ambulances incur the first two components of the response delay along with a potentially different (from CFR) delay from dispatch to the time they begin traveling. As in many standard models of ambulance operations, e.g., Daskin (1983), we assume that each ambulance is busy with probability $q \in (0,1)$, and that ambulances are busy or not independently of one another. We also assume that ambulance locations and availabilities are independent of those for CFRs. For any given call location x, let $y(x,t)$ denote the number of ambulances that can reach the location x within t minutes of the incident when stationed at their base.

The response time, including response delay, of the closest responder, whether CFR or ambulance, to a patient at Location x is a random variable, which we denote $T(x)$. The number of CFRs that are available within t minutes from a patient at Location x is, due to the Poisson point process assumption, Poisson

distributed with mean $\mu(B(x,t)) = n\alpha\nu(B(x,t))$. As is standard for Poisson point processes, the probability $\mathbb{P}(T(x) > t)$ that the response time of the closest responder is greater than t is then

$$\mathbb{P}(T(x) > t) = q^{y(x,t)} \exp\left(-\mu(B(x,t))\right). \tag{2.1}$$

Here we used the fact that the response time is greater than t if and only if all ambulances stationed at nearby bases are busy and there are no CFRs in the set $B(x,t)$. The number of CFRs in the set $B(x,t)$ is a Poisson distributed random variable, Z say, with mean $\gamma = \mu(B(x,t))$, and for such random variables, $\mathbb{P}(Z = 0) = e^{-\gamma}$.

To illustrate how one might perform these calculations, suppose a CFR walks with a constant pace of w km/min, and let d_t be the distance in km for which the response delay plus walking time equals t, for $t \geq \kappa$. We can convert response times $t \geq \kappa$ to distances using $d_t = w(t - \kappa)$. We further assume that t is so small that the CFR density may be assumed to be constant throughout the ball $B(x,t)$. Letting ω denote the value of the assumed-constant available-CFR density in the ball $B(x,t)$, we get

$$\mathbb{P}(T(x) \leq t) = 1 - q^{y(x,t)} \exp\left(-\mu(B(x,t))\right) = 1 - q^{y(x,t)} \exp\left(-\omega\pi d_t^2\right). \tag{2.2}$$

The number of sufficiently close ambulances $y(x,t)$ is readily computed for any fixed x and t given deterministic travel time models. Appendix A.2 shows the CFR response time distributions (not counting ambulances) that follow from our model.

Our discussion thus far assumes a fixed call location x. The unconditional response time, T say, is a mixture of the call-location specific response-time random

variables. Assume that calls arise over the city according to a probability distribution Λ, where $\Lambda(B)$ is the probability that a given call arises within the ball B. Then, $\mathbb{P}(T \leq t) = \int \mathbb{P}(T(x) \leq t)\Lambda(dx)$.

For computational convenience, we subdivide the city into a finite number of regions indexed by $l \in \mathcal{L} = \{1, 2, \ldots, \ell\}$. The location of a generic call lies in Region l with probability λ_l, $l \in \mathcal{L}$, so that $\sum_{l\in\mathcal{L}} \lambda_l = 1$. For all call locations x within Region l, we assume that response times $T(x)$ are identically distributed as in (2.2), with distribution function $P(T(x) \leq t) = 1 - q^{z(l,t)} \exp(-\mu_l \pi d_t^2)$. Here, $z(l,t)$ is the location-discrete equivalent of $y(x,t)$ and yields a region-specific number of nearby ambulances and μ_l is a region-specific CFR density. Thus, $\mathbb{P}(T \leq t) = 1 - \sum_{l\in\mathcal{L}} \lambda_l q^{z(l,t)} \exp(-\mu_l \pi d_t^2)$.

This model for the CFR response assumes that the density of CFRs and the number of nearby ambulances are constant within each region $l \in \mathcal{L}$, and that the response-time distribution does not vary as the call location varies within the region. This is an imperfection in the model, since calls that arise close to the boundary of a region may receive a CFR response from a neighboring region where the CFR density is different, but this effect is not captured. This model is therefore plausible when regions are large enough that CFRs outside the region are unlikely to materially impact survival rates within the region, or in the situation when neighboring regions have similar CFR densities. Appendix A.3 numerically checks the effect of the constant density assumption on our case study and concludes that, overall, it is minor. A similar discretization effect occurs for the modelling of the ambulance response, in line with existing literature modelling ambulance response.

2.4 The Ideal CFR Distribution

In view of (2.1), it is clear that for any fixed ball B we want $\mu(B)$ to be as large as possible. Recall that $\mu(B) = n\alpha\nu(B)$, where n is the number of CFRs, α is the probability that a CFR is available, and $\nu(B)$ is the conditional probability that a given CFR can be found in the ball B given that they are available. Accordingly, to make $\mu(B)$ large we can increase the number of CFRs n, we can increase the probability α that a given CFR is available, or we can influence the probability distribution ν that describes where a given available CFR is found across the city. We assume that α is exogenous to our model, though reality is potentially more complicated.

The probability distribution ν is not under our direct control. Nevertheless, it is worth identifying the *choice* of probability distribution ν that is most beneficial for two reasons:

1. An optimal CFR distribution can guide efforts in recruiting new volunteers, by indicating populations of potential CFRs that are likely to have a large impact.
2. The performance of an optimal CFR distribution can be used to determine a lower bound on the number n of volunteers needed to reach response-time goals.

With these reasons in mind, we now identify the optimal probability distribution ν for a given demand distribution λ.

First, we investigate the ideal CFR distribution across the city, where regions in the city are linked only through a bound on the total CFR mass to be allocated.

Then, we incorporate the reality that an individual recruited from Region l has a region-specific *profile*, i.e., a contribution of their CFR mass to all regions across the city.

As in Section 2.3, we assume that the locations in a city are partitioned into a finite list of neighborhoods or regions indexed by $l \in \mathcal{L}$ as above, with λ_l denoting the conditional probability that a call originates within Region l, given that a call arises. Rather than optimizing over all possible probability measures ν, which is an infinite-dimensional optimization problem, we restrict attention to probability distributions ν that are uniform within a region, so they can be parameterized by the probability mass ν_l associated with Region l, with $\sum_{l \in \mathcal{L}} \nu_l = 1$. This greatly reduces the complexity of determining an optimal probability distribution, since the calculation becomes finite dimensional. This is also reasonable in practice, since our goals relate to the overall distribution of CFRs across a city and not to the small-scale detail of exactly how they are distributed within small areas. To quantify the impact of dividing CFRs into a finite set of homogeneous regions, we compare our results with those of a more common approach: a very fine-grained but discrete network. This analysis in Appendix A.3 confirms that our approach is reasonable.

We consider the contribution of CFRs to both 1) coverage with respect to a fixed response-time threshold, and 2) a function of the full response-time distribution. The first is a standard objective in EMS; the second is more general and can capture, e.g., the average response time and the probability of patient survival (under the assumption that it is known how survival declines with response time). The details for the first objective are more straightforward so we begin there. Moreover, insights obtained from the structure of the solution based on the first

objective can help explain the solutions with respect to more complex objectives. Henceforth, when we discuss or illustrate the second type of objective, we will phrase it as 'survival', although those analyses have a broader applicability and can be performed for essentially any monotone function of the response time, including the mean.

2.4.1 Optimizing Coverage

Coverage is defined as the fraction of demand that can be served within a predefined threshold time τ. For our purpose, it is slightly more convenient to work with the probability that the response time *exceeds* τ, conditioning on the Region l in which the call occurs and using (2.2). Let ν_l denote the CFR probability mass within Region l and let a_l be the area of Region l in square km. Thus, within Region l, the density of available CFRs is $n\alpha\nu_l/a_l$. Recalling that d_τ is the distance within which a CFR can reach a call within the time threshold τ, and $z(l,\tau)$ is the number of ambulances that can reach a call in Region l within time τ, we get

$$\mathbb{P}(T > \tau) = \sum_l \lambda_l \mathbb{P}(T(l) > \tau) = \sum_{l\in\mathcal{L}} \lambda_l q^{z(l,\tau)} \exp\left(-\pi d_\tau^2 n\alpha\nu_l/a_l\right). \tag{2.3}$$

The expression (2.3) is (jointly) convex in the probabilities $(\nu_l : l \in \mathcal{L})$, yielding

Proposition 2.4.1. *The probability that the response time is greater than any fixed quantity τ is a convex function of the probabilities $(\nu_l : l \in \mathcal{L})$.*

A consequence of Proposition 2.4.1 is that we can use convex optimization methods to minimize the probability that the response time is greater than the time threshold τ, $\mathbb{P}(T > \tau)$, subject to the simplicial constraints that $\sum_{l\in\mathcal{L}} \nu_l = 1$

and $\nu_l \geq 0$ for all $l \in \mathcal{L}$. For notational simplicity let the *residual call probability* $\tilde{\lambda}_l = \lambda_l q^{z(l,\tau)}$, which is the probability a call arises in Region l and all nearby ambulances are busy, and let $\theta_l = \pi d_\tau^2 n\alpha/a_l$, for $l \in \mathcal{L}$. The optimization problem we want to solve is to

$$\begin{aligned} \min_{\nu} \quad & \sum_{l\in\mathcal{L}} \tilde{\lambda}_l e^{-\theta_l \nu_l}, \\ \text{s/t} \quad & \sum_{l\in\mathcal{L}} \nu_l = 1, \\ & \nu_l \geq 0, l \in \mathcal{L}. \end{aligned}$$

The objective function is separable, and each term is convex and decreasing in ν_l. Accordingly, a standard application of the Karush-Kuhn-Tucker conditions ensures that a greedy allocation is optimal. To describe this allocation, let us call $\tilde{\lambda}_l \exp(-\theta_l \nu_l)$ the *late rate from Region* l. We define the *marginal benefit* of adding CFR mass to Region l to be the absolute value of the derivative of the late rate with respect to ν_l, namely $\tilde{\lambda}_l \theta_l e^{-\theta_l \nu_l}$. Starting from $\nu_i = 0$ for all i, we begin by increasing CFR mass across all regions l in the set, I say, of regions having maximal marginal benefit. In doing so, the marginal increase in CFR mass in Region l should be inversely proportional to θ_l, and, in particular, proportional to the area a_l, so that the marginal benefit remains equal across all regions $l \in I$. In other words, the CFR *density* in all regions $l \in I$ increases at the same rate. This continues until either the total CFR mass 1 is used up, or the maximal marginal benefit is reduced to the point where a new Region j joins the set I of regions attaining the maximal marginal benefit. A more precise description of the algorithm can be found in Appendix A.4.

The value of ν_l for each $l \in I$ at which a new region joins the index set I can be computed in closed form. Let $j \in \arg\max\{\tilde{\lambda}_i \theta_i : i \notin I\}$ be the maximal marginal

benefit of any region that is yet to receive positive CFR mass, and let $l \in I$ be arbitrary. We solve $\tilde{\lambda}_l \theta_l e^{-\theta_l \nu_l} = \tilde{\lambda}_j \theta_j$. Using the definition of θ_l and θ_j, cancelling common terms, and solving for ν_l yields

$$\nu_l = \frac{1}{-\theta_l} \ln\left(\frac{\tilde{\lambda}_j / a_j}{\tilde{\lambda}_l / a_l}\right) = \frac{a_l}{\pi d_\tau^2 n \alpha} \ln\left(\frac{\tilde{\lambda}_l / a_l}{\tilde{\lambda}_j / a_j}\right). \tag{2.4}$$

A proviso is that we do not reach this new value if we "hit" the constraint $\sum_{l \in I} \nu_l = 1$ first. In this case, we do not reach the point where ν_j becomes positive.

The expression (2.4) yields several insights that are reinforced in a two-region example in Appendix A.5. First, regions are not guaranteed to receive positive probability mass; excluded regions are those with the lowest residual call densities $\tilde{\lambda}_l / a_l$. Second, regions receiving positive probability mass do so roughly in proportion to the product of their area and the log of their residual call densities. Since the density of CFRs is proportional to ν_l / a_l, and thus divided by area, it follows that regions receiving positive CFR density do so roughly in proportion to the log of their residual call densities. We use the term "roughly" because (2.4) is not the precise form of the optimal allocations, as mentioned above. Third, the expected number of CFRs allocated to a Region l that receives positive mass is roughly $n\nu_l$, which, from (2.4) does not depend on n since ν_l is proportional to n^{-1}, where "roughly" has the same sense as before. Finally, the optimal allocations are roughly proportional to the inverse of the area, πd_τ^2, of the circle within which CFRs can respond within time τ. For smaller τ (more stringent targets) the optimal allocations are thus more concentrated in high-call-density regions.

2.4.2 Optimizing Survival

The same approach works for more complex performance metrics, which we illustrate by showing that maximizing the probability of survival for OHCA is a convex optimization problem in the CFR location probabilities (ν_l, $l \in \mathcal{L}$).

Proposition 2.4.2. *If the probability of death is increasing in the response time, then the probability of survival is a concave function of the probabilities* $(\nu_l : l \in \mathcal{L})$.

Proof. We show that the probability of death is convex in the CFR location probabilities (ν_l, $l \in \mathcal{L}$). We assumed that the death probability is an increasing function, $\bar{f}(\cdot)$ say, of the response time T, which includes the response delay κ. Increasing functions have at most a countable number of discontinuities, so we can adjust $\bar{f}$ to an increasing, right-continuous function, f by redefining $\bar{f}$ in at most a countable number of points. Thus, we now work with the increasing, right-continuous function f. Let $\underline{d} = f(\kappa) \in [0,1]$ be the minimal death probability assuming immediate ambulance response. Let $\bar{d} \in [0,1]$ be the maximal death probability assuming response time is infinite. Using the fact that the mean of a nonnegative random variable Y can be expressed as $\int_0^\infty \mathbb{P}(Y > u)\, du$, the death probability as a function of ν, $d(\nu)$ say, is

$$d(\nu) = \mathbb{E}f(T) = \int_0^1 \mathbb{P}(f(T) > u)\, du = \underline{d} + \int_{\underline{d}}^{\bar{d}} \mathbb{P}(T > f^{-1}(u))\, du,$$

where $f^{-1}(u) = \inf\{r : f(r) \geq u\}$ for $u \in [0,1]$ and the final step follows from Serfling (1980, Lemma 1.1.4(iii)). Hence, conditioning on the Region l of the call,

$$\begin{aligned} d(\nu) &= \underline{d} + \sum_{l \in \mathcal{L}} \lambda_l \int_{\underline{d}}^{\bar{d}} \mathbb{P}(T(l) > f^{-1}(u))\, du \\ &= \underline{d} + \sum_{l \in \mathcal{L}} \lambda_l \int_{\underline{d}}^{\bar{d}} q^{z(l, f^{-1}(u))} \exp\left\{-|B(l, f^{-1}(u))| n\alpha\nu_l / a_l\right\} du. \end{aligned} \tag{2.5}$$

The integrand in (2.5) is convex in ν_l. □

Thus we can efficiently maximize the survival probability over probability distributions ν on $\mathcal{L}$. In fact, a greedy approach that adds CFRs to regions where they yield the highest marginal reduction in death rate gives an optimal solution. In contrast to the greedy approach for late arrivals, the marginal allocation over the different regions does not remain constant. Therefore, an expression similar to (2.4) to compute the *exact* CFR distribution does not hold. Instead, we use a discretized step size ϵ, which is optimal as $\epsilon \to 0$. We implemented this method and give numerical results in Section 2.6.

2.4.3 More Complex Survival Functions for OHCA

So far we have focused on survival functions that depend only on the response time, whether by CFR or ambulance. More complex survival functions for OHCA patients such as 'Model 1' in Waalewijn et al. (2001) depend on both the time to CPR (t_{CPR}) and the time to EMS arrival (t_{EMS}), where $t_{\text{CPR}} = t_{\text{EMS}}$ when an ambulance arrives before volunteers, and t_{CPR} equals the response time of the volunteer otherwise. Let $f(t_{\text{CPR}}, t_{\text{EMS}})$ be the probability of death as a function of the time to CPR and the time to EMS arrival. Model 1 of Waalewijn et al. (2001) uses

$$f(t_{\text{CPR}}, t_{\text{EMS}}) = 1 - (1 + e^{0.04+0.3t_{\text{CPR}}+0.14(t_{\text{EMS}}-t_{\text{CPR}})})^{-1}. \tag{2.6}$$

These times are measured from patient collapse instead of call initiation. Waalewijn et al. (2001) estimates the difference between these two moments to be 1 minute (median). Recall that $z(l, \tau)$ is the number of ambulances that can reach Region l from their home base within τ minutes of cardiac arrest. As before, let $d(\nu)$ be the probability of death for a single call as a function of the CFR location distribution ν. Let a_{jl} denote the response time of the jth closest ambulance

to Region l when all ambulances are available and at their respective bases. Let $T_v(l)$ be the (random) response time of the closest volunteer, the distribution of which can be found by dropping the $q^{y(x,t)}$ term in (2.2). Conditioning on the location of the call and the closest available ambulance to Region l, we get

$$\begin{aligned} d(\nu) &= \sum_{l\in\mathcal{L}} \lambda_l \mathbb{E} f(T_{\text{CPR}}(l), T_{\text{EMS}}(l)) \\ &= \sum_{l\in\mathcal{L}} \lambda_l \left(q^k \mathbb{E} f(T_v(l), \infty) + \sum_{j=1}^{k} q^{j-1}(1-q) \mathbb{E} f(\min\{T_v(l), a_{jl}\}, a_{jl}) \right). \quad (2.7) \end{aligned}$$

The term $f(T_v(l), \infty)$ corresponding to the situation where all ambulances are busy is 0 according to Model 1 of Waalewijn et al. (2001), but more generally is convex in ν so long as f is increasing in its first argument, as we have already established. As to $\mathbb{E} f(\min\{T_v(l), a_{jl}\}, a_{jl})$, so long as $f(t_{\text{CPR}}, t_{\text{EMS}})$ is increasing in t_{CPR} for any fixed value of t_{EMS}, which we henceforth assume, then our previous arguments extend to ensure convexity in ν. Thus, we have established

Proposition 2.4.3. *Suppose that $f(t_{CPR}, t_{EMS})$ is increasing in t_{CPR} for each fixed $t_{EMS} \in [0, \infty]$. Then the probability of survival is a concave function of the probabilities $(\nu_l : l \in \mathcal{L})$, even when accounting for the impact of ambulances providing advanced care in addition to CPR.*

2.4.4 Profile Optimization

Till now we have assumed that CFRs spend all their time in a single region, yet in reality a person likely spends their time in a few different regions during different times of day. So, instead of having one specific location probability distribution ν, we have time-dependent CFR location distributions $\nu^{(t)}$, where t is assumed to fall into one of a small number of discrete time blocks $\mathcal{T}$, e.g., $\mathcal{T} = \{$weekday,

weeknight, weekend day, weekend night}. A CFR provides a contribution in each time $t \in \mathcal{T}$, and for any fixed t can contribute to several elements of $\nu^{(t)}$. We model these contributions through *time-dependent profiles*, in which people with the same profile are assumed to spend their time in the same way - at least probabilistically speaking. That is, we focus on the aggregate contribution of all CFRs of a certain profile.

Assume CFRs take one of m different profiles. For each time block $t \in \mathcal{T}$, let $\mathcal{P}^{(t)} \in [0,1]^{|\mathcal{L}|\times m}$ denote the given CFR profile matrix, in which $\mathcal{P}_{li}^{(t)}$ is the probability, in Time Block t, that a CFR with Profile i appears in Region l. We have $\sum_{l\in\mathcal{L}} \mathcal{P}_{li}^{(t)} \leq 1$ for all $t \in \mathcal{T}$, with strict inequality when CFRs temporarily leave the city.

Let x_i be a decision variable denoting the proportion of CFRs recruited with Profile i for $i = 1, 2, \ldots, m$ such that $\sum_{i=1}^{m} x_i = 1, x_i \geq 0$. The resulting CFR location distribution per time block can then be calculated as

$$\nu^{(t)} = \mathcal{P}^{(t)} x, \quad t \in \mathcal{T}. \tag{2.8}$$

Our goal is to optimize either coverage or patient survival obtained from this $\nu^{(t)}$ with decision variables $(x_i \colon i = 1, 2, \ldots, m)$ and constraints (2.8) ensuring the link between recruitment and CFR location.

Proposition 2.4.4. *For both the coverage and survival objectives, the optimization problem with profiles remains a convex optimization problem under the same conditions stated in Proposition 2.4.2.*

To prove Proposition 2.4.4, observe that the objective function becomes a weighted sum of (2.3) or (2.5) over the set of time blocks $\mathcal{T}$, which remains convex.

Moreover, the constraints we add are linear. Hence the problem remains convex under both objectives.

With profiles, the optimization problems are no longer separable so a greedy approach is no longer optimal. To solve the optimization problem with profiles, we implemented a version of the away-step Frank-Wolfe algorithm in Bomze et al. (2020) with an adaptive line search and a stopping criterion based on the (optimality) gap between the objective value of the current iterate and a lower bound obtained through a cutting plane method that relies on convexity of the objective function. See Appendix A.6 for more details and the pseudo-code of the implemented Frank-Wolfe algorithm.

2.5 Case Description

Auckland is the largest city in New Zealand, with the greater Auckland region having a population of 1.4 million and an area of 5000 square kilometers. We focus on the most urban area by excluding 7 out of 21 local boards with a population density of less than 1000 people per square kilometer. The resulting area has a population of 1.1 million, covers 500 square kilometers and is divided into 287 so-called area units, yielding the set of regions $\mathcal{L}$. These area units are the second smallest unit at which Statistics New Zealand collects data (Stats New Zealand Geographic Data Service 2016).

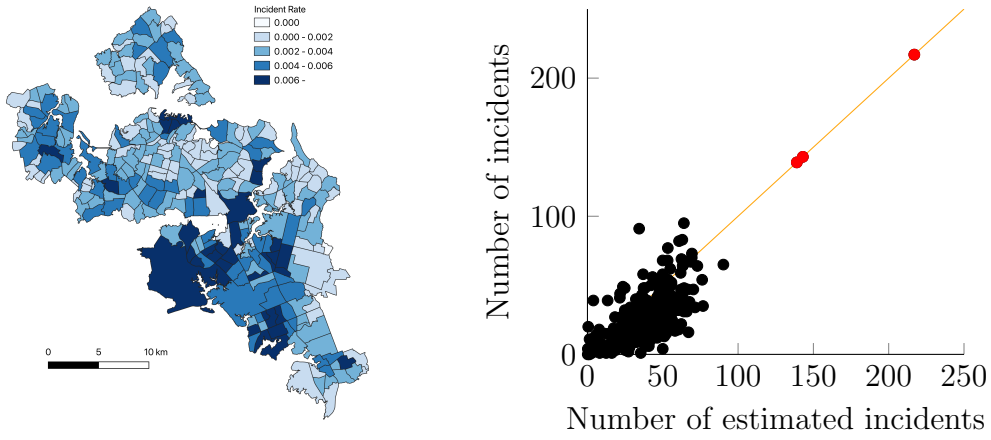
2.5.1 OHCA Incidence Rates

The SJAS has provided us with the time and location of all OHCAs in Auckland since 2013. One could use the empirical OHCA rate to estimate the demand vector λ, but low counts in some area units have very high variability relative to their mean. Instead, we estimate the OHCA incident rate by combining demographics with socio-economic factors as in Dicker et al. (2019).

Figure 2.1 shows both the empirical incident rate of different area units and a scatter-plot of the number of OHCA incidents since 2013 versus the number of estimated OHCA incidents using Dicker et al. (2019), where each point represents one area unit. We see a general linear trend as expected, but the OHCA incident counts are dispersed relative to a linear trend. The three area units highlighted in Figure 2.1b are two area units covering the Central Business District (CBD) and the area unit containing the airport. We replaced the estimated rates in these 3 area units with the empirical counts in SJAS data, because these area units were extreme outliers in the original fit. The input demand rate λ to our model is then the normalized estimated OHCA counts after the CBD and airport adjustments. More details and discussion on this estimation process can be found in Appendix A.7.

2.5.2 Response Process

We assume that an ambulance always responds and that volunteers are alerted if they are close to the incident. Of interest is the time it takes for the first responder to arrive on-site. All reported response times are measured from call initiation, meaning that they do not include the time between cardiac arrest and



(a) Empirical Incident Rate (b) Empirical vs Estimated Incidents

Figure 2.1: Left: The empirical incident rate per area unit in Auckland, New Zealand. Right: Scatter plot of the actual and estimated number of cardiac arrests between 2013 and 2020 over all area units.

call initiation. This is in line with classical performance measures for EMS systems and is the only time known to dispatchers. A response time, whether by ambulance or by volunteer, consists of a response delay and a travel time. We base our estimates on GoodSAM data whenever available and reliable; in other cases, we resort to estimates in the literature.

CFR Response

The local CFR app GoodSAM is run by the St. John Ambulance Service (SJAS). We have obtained data from the SJAS on CFR operations since the inception of the program in December of 2017. Between December 2017 and November 2020, 6749 CFR alerts were sent out for 2827 OHCA incidents. We provide an exploratory data analysis in Online Appendix A.8 that suggests that our model is reasonably consistent with the very noisy data at hand.

For the probability that a volunteer accepts an alert, we use the historical value

as observed in the data: $\alpha = 0.14$. We take this value to be equal for all CFRs rather than assuming an individual-specific acceptance probability because our data is insufficient to accurately estimate such values. We assume that accepting CFRs respond on foot, and use the most conservative estimate of Slaa (2020) in the Netherlands to model their speed at 6 km/h. We assume a constant response delay, which consists of a triage and system activation time as well as a delay for CFR acceptance. For triage and system activation we assume 1.5 minutes; this is in line with Slaa (2020) who estimated this time at 105 seconds before a system update that automates and thereby speeds up system activation. For the time it takes for a volunteer to accept or reject (CFR acceptance delay), we take the mode of the GoodSAM data: 30 seconds. The sum gives a response delay of 2 minutes. We further assume that only CFRs within 1 km of the patient are dispatched, in line with current practice. This implies that the response time of CFRs, conditioned on finding at least one available CFR within 1 km, is between 2 and 12 minutes.

The assumed 30 second CFR acceptance delay is not perfect, because volunteers may accept an alert but not initiate their departure immediately, or they may accept the alert after they start traveling. To see the impact of a shorter or longer acceptance delay, one can refer to adjacent columns in Table 2.1.

Ambulance Response

The SJAS operates 15 bases from which they could realistically respond to the part of the city that we analyze. We model the ambulance pre-trip delay as 2 minutes; Ridler et al. (2022) estimates this value to be 2.1 minutes in Auckland. We used driving times for EMS vehicles between each base and the centroid of each area unit as obtained from an open-source ambulance-simulation package (Ridler 2020);

see also Ridler et al. (2017, 2022).

Since OHCA patients only constitute a small portion of all the requests served by SJAS, it is not reasonable to assume the ambulances are placed optimally with respect to OHCA response targets. Instead, we assume that ambulances are placed to meet the SJAS target of responding within 12 minutes for at least 95% of their patients (Ministry of Health 2020). We obtained the corresponding ambulance locations as follows. We assume that each ambulance has a utilization of 0.44, as used in Ridler et al. (2022) and distribute ambulances across the existing bases according to the optimal solution of MEXCLP (Daskin 1983). We use 25 ambulances, which is just sufficient to achieve the mentioned target (95% within 12 minutes, including a pre-trip delay of 2 minutes).

2.5.3 CFR Profiles

For the optimization including CFR movement as introduced in Section 2.4.4, we need data on profiles. These profiles within a given time block can be interpreted as giving the probability that a certain CFR appears in a given area unit in that time block. We introduce a profile for each area unit and assign potential CFRs a profile based on their home address. We distinguish two time blocks: working hours (weekdays from 8 AM till 10 PM) and all hours outside of that, which we will refer to as 'nights and weekends'. Thus, the interpretation of $\mathcal{P}_{li}^{(t)}$ is the probability that a CFR living in area unit i appears in area unit l in time block t. Although we defined profiles per area unit, it is also possible to have more - or fewer- profiles than area units, for example based on other location aggregation levels, e.g., people who work at the airport, or other properties on which one can distinguish recruitment, e.g., students.

To obtain profile data for working hours, we combine the population of age 15 years or older from the 2013 census (Stats New Zealand Geographic Data Service 2015a) with the commuting routes for the resident population (Stats New Zealand Geographic Data Service 2020). These data give the work location of residents from each area unit in New Zealand. We then assume that working potential CFRs spend their time in the area unit of their work location and non-working potential CFRs spend their time in the area unit in which they live. The resulting profiles have on average 62% of people on the diagonal, so a CFR recruited from a certain area unit will, on average, contribute to another part of the city for 38% of the 'working hours' time block. For the profiles during nights and weekends, we assume that everyone stays in the area unit where they live.

This fitting process is not perfect, but we consider the approach sufficiently realistic to illustrate our models. Indeed, a closely related approach was used in Cont et al. (2021) to model population mobility in a study of intervention policies during the COVID-19 pandemic. Alternative sources were infeasible for this case study because they either were not sufficiently fine-grained (Facebook Data For Good 2021) or were not available for Auckland (Safegraph 2021). We discuss profile estimation impact on performance and provide a sensitivity analysis in Appendix A.9.

2.5.4 Performance Metrics

To measure late arrivals, we consider the two response-time targets under which the SJAS operates: 95% of their incidents should be reached within 12 minutes and 50% should be reached within 6 minutes (Ministry of Health 2020). The other metric we consider is patient survival, which we quantify using (2.6). This

function was fitted on data where both ambulances and volunteers responded, and volunteers did not carry a defibrillator. This function takes as input t_{CPR} and t_{EMS} which are measured from patient collapse, unlike response time which is measured from call initiation. In line with Waalewijn et al. (2001)'s estimate, we add 1 minute (median time to call initiation) to account for this difference.

2.6 Numerical Results

In this section we answer the four questions posed in the introduction: 1) In what areas is introducing a CFR app an efficient way to reach certain response-time targets? (Section 2.6.1); 2) What is the benefit of introducing a CFR app for a base of already trained volunteers, with known home addresses? (Section 2.6.2); 3) What is the benefit of a CFR app with a given number of registered volunteers, when you do not know where they live or spend their time? (Section 2.6.3); and 4) Where to recruit additional CFRs for an existing app with known current CFR base? (Section 2.6.4). The first two questions are answered using the models from Section 2.3; the last two questions relate to the optimization models presented in Section 2.4. For clarity, ambulances are omitted in Section 2.6.1. The results in all other sections incorporate ambulance response as outlined in Section 2.5.2. We illustrate our calculations with data from the case study described in Section 2.5.

2.6.1 CFR Density Requirements

We next analyze the response-time distributions derived in Equation (2.2) and for clarity, omit ambulances. The CFR response-time distributions are plotted

in Figure A.1. These allow us to calculate the required CFR density to reach a given response-time target in closed form. This can be used to investigate in what areas introducing a CFR app would be an efficient way to meet a certain target. Table 2.1 gives the required density of CFRs to meet a target that is defined by a combination of a time limit and a compliance level.

Table 2.1: Required density of available CFRs (per km^2) to meet a range of response-time targets, assuming a 2-minute pre-trip delay and a walking speed of 6 km/h.

		Response-time target (minutes)									
		3	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Cov.	**0.5**	22.06	5.52	2.45	1.38	0.88	0.61	0.45	0.34	0.27	0.22
	0.7	38.32	9.58	4.26	2.40	1.53	1.06	0.78	0.60	0.47	0.38
	0.9	73.29	18.32	8.14	4.58	2.93	2.04	1.50	1.15	0.90	0.73

The results presented in Table 2.1 are generic in the sense that they apply to any location. They assume a fixed walking speed and pre-trip delay. The table gives the required density of *available* CFRs; the required number of *registered* CFRs can be obtained by dividing the given density by the acceptance probability α.

We illustrate the application of these results for the city of Auckland, where the observed value of α is 0.14. For each area unit, we can compute the fraction of the area-unit population that needs to be registered in order to satisfy the response-time requirements. We take both requirements (90% in 12 minutes and 50% in 6 minutes) and observe that the latter requires more CFRs: reaching this target requires a density of $1.38/0.14 = 9.19$ registered CFRs per square kilometer.

Looking at the population density of Auckland, we see that this density translates to 0.42% of the population registered as CFRs. However, the population

density fluctuates over the city so we need different registration rates in different parts of the city. A uniform registration rate of 0.42% of the population would meet the target only in 217 out of 287 area units. The required registration rate per area unit fluctuates between 0.08% and 130% of the population. Under the assumption that registration rates will not exceed 1% (a number currently achieved in countries with high participation), at least 23 area units (covering 3.7% of the population) will not be able to meet the specified target. Thus, even in an urban area like Auckland, the population density is not sufficiently consistent to obtain good response times throughout. Given that organizations are evaluated based on aggregate performance, one might observe different CFR densities in different area units that together do achieve an overall response-time requirement, which we explore next.

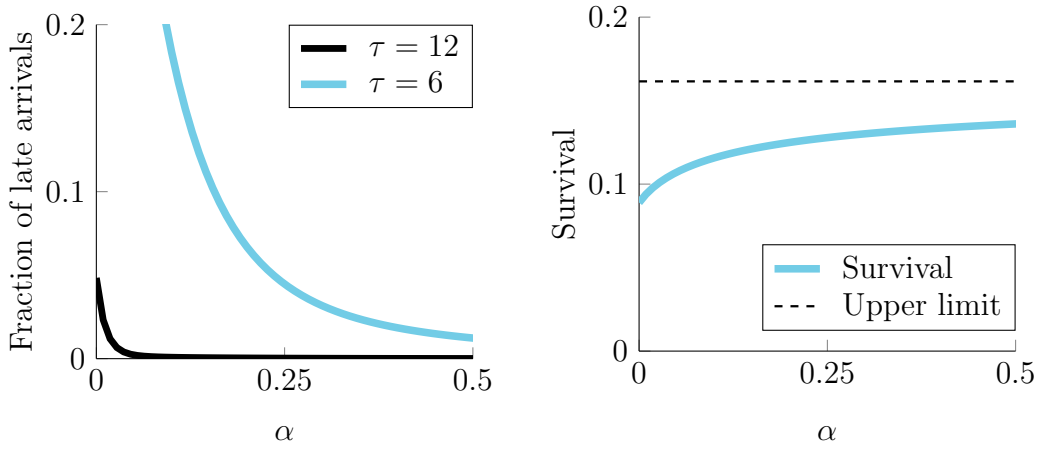
2.6.2 Quantifying the Benefit of a Given CFR Base

Next, we quantify the benefit of introducing an alert app, in a scenario with a known base of already trained CFRs and a given ambulance distribution. As already noted, we have been contacted by a party facing exactly this question. They have a base of trained CFRs that is normally used at events such as festivals or parades, and are considering introducing an alert system to expand CFR impact.

When we say a given CFR base is known, we mean that n and ν are known, e.g., because CFRs registered their home and/or work address or because they allow GPS tracking. We quantify how such a known CFR base would contribute to the fraction of late arrivals and the survival probability, and we explore how this depends on the acceptance rate $\alpha \in [0, 0.5]$. We vary α because it is, to some extent, controllable, and there is uncertainty about its true value. This range

covers values that seem plausible given Auckland data and conversations with app developers.

As the party that raised this question wants to remain anonymous, we illustrate our calculations for Auckland. We consider the case where 1% of the population of Auckland are trained CFRs and a probability measure ν that is proportional to the estimated demand. In contrast to the results from Table 2.1, for which it was sufficient to consider every area unit separately, this question requires us to look at the entire city with varying CFR and incident densities, aggregating results over the city of Auckland.



(a) Late arrivals

(b) Survival

Figure 2.2: Impact of increasing the CFR availability α on fraction of late arrivals and survival, assuming 1% of the population is signed up as CFRs, distributed over the area units proportionally to demand.

Figure 2.2a shows that bringing down response times that exceed the 12-minute target from 4.8% (without using CFRs) to 1% requires an α of 9% (1000 available CFRs). While not depicted in the figure, to satisfy the requirement of reaching 50% of the incidents within 6 minutes, $\alpha \approx 2.7\%$ suffices. In Figure 2.2b, we see that even for low values of α, CFRs contribute substantially to survival. Convergence to the upper limit of survival is very slow, and cannot be achieved by only increasing

α. Additional recruitment would be needed, which we discuss later.

2.6.3 Quantifying the Impact of CFRs' Geographical Spread

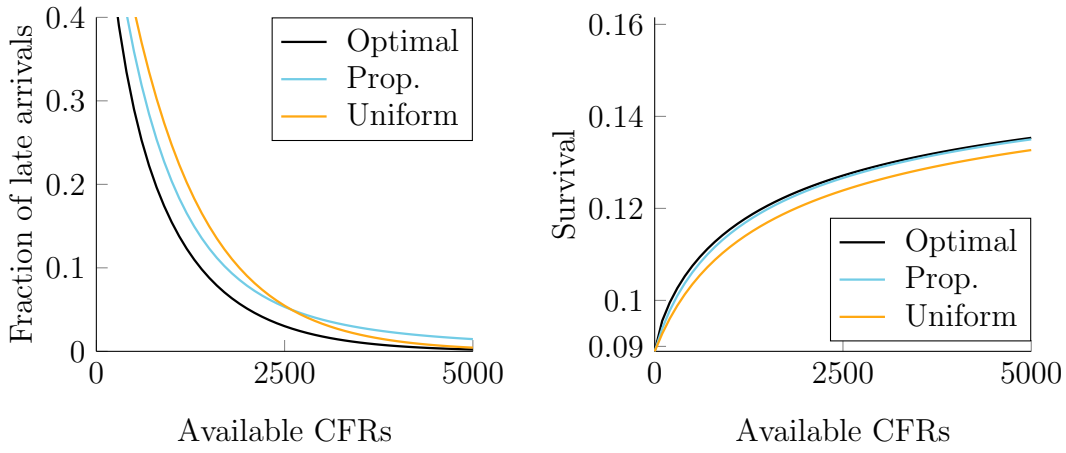
In the previous section we assumed that the CFR distribution ν was known; in contrast, we now investigate the scenario in which it is unknown where CFRs spend their time. Practitioners in Auckland are facing this scenario and would like to quantify the benefit of their app alongside the existing ambulance response. While a retrospective study is possible, this is difficult for recently introduced CFR systems, as they have limited data available and accurately estimating the response-time distribution requires many data points per area unit.

First, we consider two scenarios: one where CFRs are uniformly distributed over the city and another where CFRs are distributed proportionally to demand. The uniform distribution serves as a useful reference, while the proportional distribution is plausible. Second, we bound the performance of a CFR base of given size, using the optimization models from Section 2.4. These result in a lower bound for the fraction of late arrivals and an upper bound for the survival probability with a given total number of CFRs.

For a city that operates a CFR app, the values of n and α are likely known. However, to provide more insight, we perform this case study for different values of $n\alpha$, the *expected number of available CFRs*. As we shall see, different values result in different optimal CFR allocations to area units, requiring us to solve the optimization problem separately for each value of $n\alpha$. We run our calculations up to 5000 available CFRs, which, with the observed α of 0.14 corresponds to

approximately 3.2% of the overall population of 1.1 million.

Recall that for both objectives — late arrivals and patient survival — a greedy approach is optimal. When optimizing late arrivals, an exact implementation (up to numerical precision) of the greedy method exists, using Equation (2.4). In contrast, when optimizing patient survival, the greedy method requires a step size ϵ, which we set to 1/10,000. This means that we allocate the total number of CFRs in portions of size $n/10{,}000$ to the area unit where they would most improve the objective.



(a) Late arrivals $\tau = 6$ min

(b) Survival

Figure 2.3: Impact of increasing number of available CFRs ($n\alpha$) for different CFR distributions (ν) on the fraction of late arrivals ($\tau = 6$ min) and the survival probability. The black lines are each formed by the solution of 50 optimization problems.

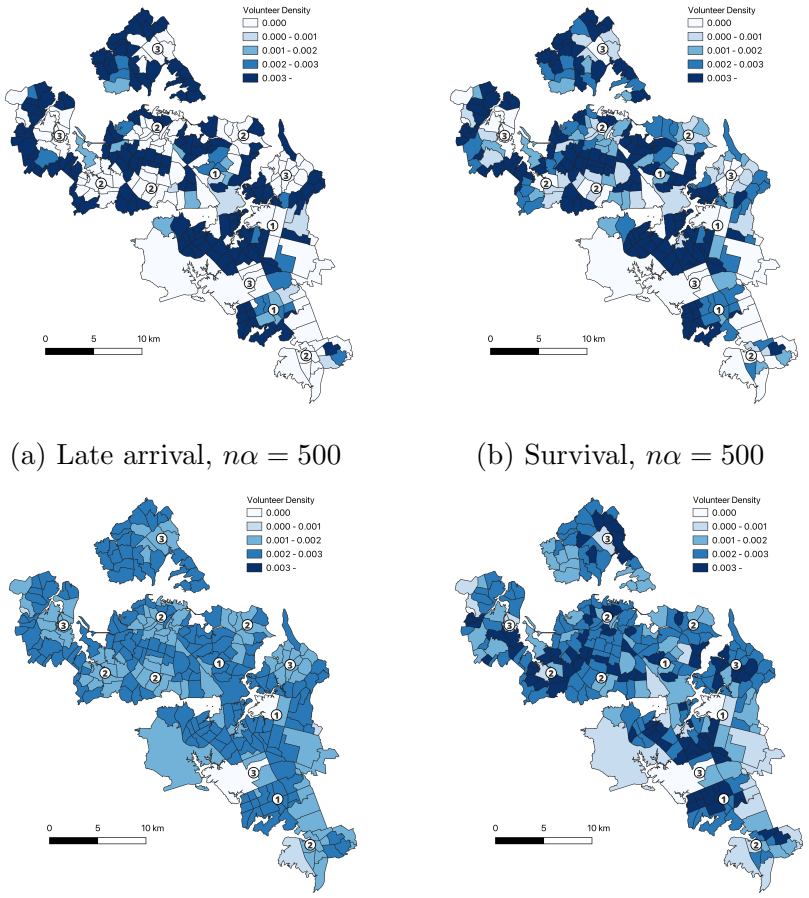
Figure 2.3 depicts the performance of the three CFR distributions, for two different objectives. Figure 2.3 shows that the impact of the system depends on the geographical distribution of the CFRs. For example, in Figure 2.3a we can see that to bring the late arrivals down to 10%, one needs either 1400 optimally distributed CFRs, 1900 uniformly distributed CFRs, or 1800 proportionally distributed CFRs. In particular, in this part of the graph, the proportional distribution outperforms

the uniform distribution.In contrast, for a large number of CFRs, the uniform distribution outperforms the proportional one and approaches the optimal performance. This observation is even more apparent in the case with a homogeneous ambulance response, for which the results are given in Appendix A.10. Not just the performance is similar, but the optimal volunteer distribution also closely resembles the uniform distribution, as illustrated in Figure 2.4. This figure compares the geographical distribution of an optimal ν for both a limited number of CFRs ($n\alpha = 500$) and a larger number of CFRs ($n\alpha = 5000$). Distinguishing two objectives (late arrivals and patient survival) yields four scenarios. In particular, Figure 2.4c shows a distribution of CFRs that resembles a uniform distribution, only distorted by the ambulances. This is in line with our earlier observations regarding the greedy algorithm. Figure 2.4a shows that for a limited number of CFRs, area units with a low call density do not receive any CFR mass.

For survival (Figures 2.4b and 2.4d), the uniform distribution does not emerge as $n\alpha$ grows; it remains optimal to focus on high-demand areas even with 5000 available CFRs, in accord with the close-to-optimal performance seen in Figure 2.3 for the proportional distribution. For a high number of CFRs, low-demand areas receive slightly more CFRs than a proportional allocation would assign to them, as is further detailed in Appendix A.11.

2.6.4 Optimal Recruitment Efforts

Next, we investigate efficient recruitment strategies while taking into account the current ambulance bases. We explore *profile recruitment*, where we take into account CFR presence across the city depending on where they live. In the previous upper bound calculations we optimized for the available CFR location distribution



(a) Late arrival, $n\alpha = 500$ (b) Survival, $n\alpha = 500$

(c) Late arrival, $n\alpha = 5000$ (d) Survival, $n\alpha = 5000$

Figure 2.4: Optimal CFR distribution (fraction of CFRs in that region divided by surface area) to minimize late arrivals for $\tau = 6$ min (a and c) or to maximize survival (b and d), with $n\alpha = 500$ or $n\alpha = 5000$. Darker colors represent higher CFR density. The white circles represent the location of the ambulance bases, and the numbers within them indicate the number of ambulances stationed at each base.

ν, which, in a scenario without profiles, is equal to x. When CFR locations are driven by profiles, the relationship between x and ν is described by Equation (2.8). Not every desired CFR probability mass ν is attainable by adjusting recruitment mass x per area unit. To find the optimal recruitment strategy, we use the formulation given in Section 2.4.4 and the method for obtaining profiles described in Section 2.5.3.

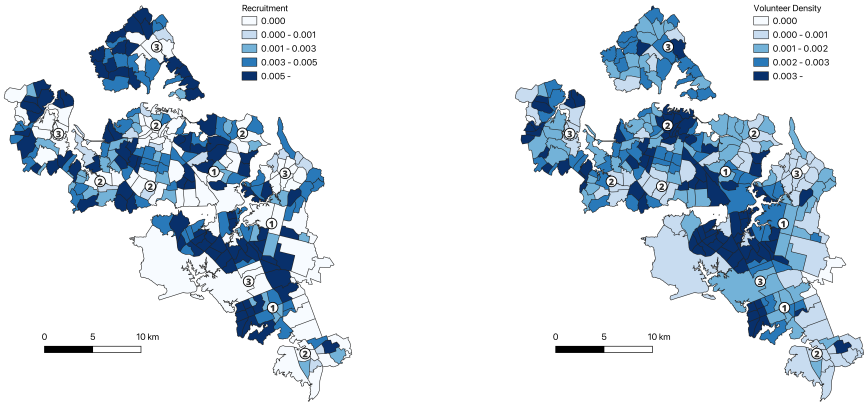
The impact of profiles on the optimal recruitment strategy can best be illustrated when we consider a homogeneous ambulance response throughout the city. We contrast the difference in recruitment per area unit with and without profiles in Appendix A.12. There we see that some area units require no recruitment under profiles because of the influx of CFRs from other area units. Some of these area units, most notably the CBD, have such a high influx of CFRs that even without recruitment from that area unit the density of CFRs exceeds the optimal density without profiles. Most other area units have slightly higher recruitment under profiles to compensate for the outflux of CFRs.

Time-dependent profiles

Thus far, we have ignored a time component so we could isolate the importance of capturing CFR movement. Next, we consider the more realistic case where profiles have both a time and a location component, as modeled in Section 2.5.3. We apply the models from Section 2.4.4 to investigate optimal recruitment under these time-dependent profiles. We illustrate our results for a survival objective and the equivalent of 500 available CFRs. Since the value of α is now time-dependent, we cannot fix the value of $n\alpha$ throughout the day but instead fix n to $n = 500/0.14 = 3571$.

Figure 2.5 shows the amount of recruitment as well as the resulting CFR density per area unit. The area around the CBD is easily recognized as an area with high influx (dark blue in Figure 2.5b), so it requires little to no recruitment. Only in the CBD itself, a minor recruitment effort is still advised, due to the disproportionately high demand there.

The solution in Figure 2.5 results in a 24-hour-average survival rate of 10.58%



(a) Optimal recruitment (x_i values)

(b) Resulting daytime density (ν_{l2})

Figure 2.5: Time-dependent profile recruitment that maximizes survival for $n = 500/0.14 = 3571$. The white circles represent the location of the ambulance bases, and the numbers within them indicate the number of ambulances stationed at each base.

city-wide, which breaks down as 10.74% for day and 10.21% at night. Optimizing for daytime only, the survival rate during the day would be 10.75%, but this improvement would not outweigh the lives sacrificed during the night.

Our case study demonstrates the possibility of capturing profiles in recruitment. Ignoring profiles altogether (both their time and location component) would lead to a minor reduction in survival: 10.56% instead of 10.58%. This difference, though small, is unnecessary. To provide some perspective, a survival difference of this magnitude is the equivalent of recruiting and training approximately 80 extra CFRs (2% of the CFR base).

2.7 Conclusion and Discussion

We introduced models to quantify the impact of CFR systems alongside the existing ambulance response. These models were subsequently used to answer four questions that are central in practice in Sections 2.6.1-2.6.4, using data from the city of Auckland, New Zealand. The results yield important insights, such as how the performance depends on the number and location of CFRs. While the results and insights are then specific to Auckland, the same models and approaches may be used to explore the potential impact of such an app in new cities, or to guide the recruitment of additional CFRs in existing systems.

Figures 2.2 and 2.3 show the steady improvement in two performance metrics (late arrival and patient survival) over a wide range of numbers of available CFRs that complement the existing ambulance contribution. The marginal benefit of additional CFRs remains practically important, even beyond optimistic registration rates ($\sim$1%). Besides increasing the number of CFRs n, one may take measures to increase the availability probability α. The importance of increasing α is underlined by the insight that doubling α has the same impact as doubling the number of CFRs. Thus, systems must be designed to maximize α.

Every additional CFR improves the system-wide response time, but where those CFRs spend their time matters. Thus, it makes sense to steer recruitment efforts to the right areas. These areas can be determined with the models from Section 2.4. An important empirical observation from Section 2.6 is that optimal CFR densities heavily depend on the objective. To maximize cardiac arrest survival, a volunteer distribution proportional to demand is nearly optimal. Conversely, for late arrivals we see clear differences between uniform volunteer densities (if the time threshold is loose) and clustered volunteers (if the time threshold is strict). Thus, the objective

heavily influences the equity of the final solution. A uniform distribution is, at least in the absence of ambulances, perhaps the most equitable. However, having equal volunteer densities everywhere would require an impractically high sign-up rate in sparsely populated areas. If equity needs to be explicitly captured, our model can be extended by specifying a lower bound on the performance in each area without overly complicating the solution method.

For OHCA, we consider maximizing survival an appropriate objective, and under that assumption, a first conclusion appears that incident rates can guide recruitment efforts. However, CFRs do not spend all of their time in the area unit where they live, and these effects typically do not even out. Consequently, even without *any* recruitment from the city center, our daytime scenario showed an over-resourcing of CFRs in the city center. We expect similar effects in other cities. The results from Section 2.6.4 show that this indeed makes a significant difference in terms of where to best recruit. If one knows how volunteers tend to move through the city, these dynamics can be incorporated in the form of profiles, and our method remains able to find the optimal recruitment. An additional use of the profile concept is to quantify the marginal benefit of training a new CFR based on that person's profile, which could inform, e.g., an organization's willingness to fund the CFR's training. In this manner, our model allows personalized advice and ensures that a limited training budget can be spent in an effective way.

As mentioned in the introduction, the practical value of our work has already been demonstrated through studies with the Fire Department Amsterdam-Amstelland in the Netherlands and a Red Cross organization (anonymous, but at a national level). Similar ideas are applicable in settings with many uncontrollable servers who can thus be modeled as randomly situated throughout an area. One

of the authors has used these ideas in that manner, but confidentiality agreements prevent the disclosure of the setting.

There are many important directions for future research. First, a detailed analysis of data from CFR schemes is needed to understand CFR responses. For example, how do acceptance probabilities depend on the distance to the patient or personal attributes of the CFR such as age or mobile phone specifications? Second, how should dispatch strategies be designed, e.g., given the distances of nearby CFRs, which should be notified in a first batch of alerts, and what is the benefit to retaining notifications for very close CFRs after a more distant CFR accepts? Should one avoid over-utilization of CFRs and navigate a trade-off between notifying CFRs for the current task and saving them for future ones? This question is posed in more detail in Henderson et al. (2022) and explored in a more general setting in Manshadi and Rodilitz (2022).

It may appear that we have ignored the impact of bystanders on survival rates. Even though they are unlikely to be trained in CPR, De Maio et al. (2003) found that most patients who survived received some sort of bystander intervention. This bystander help was already included in the patient data that was used to train survival functions. In the absence of more comprehensive data, we have tried to incorporate the bystander effect to the best of our ability by using survival functions that account for bystander involvement.

In the context of OHCA, whether or not an AED is available has a significant impact on survival. For example, what is the additional benefit of taxi drivers carrying AEDs in their cars and being dispatched to nearby incidents, even if they do not have medical training? What survival benefits can be expected if an AED gets delivered to the scene by a drone? How valuable would it be to provide a

small fraction of volunteers with their own AED? A mix of volunteers carrying AEDs supplemented with drone-delivered AEDs could by optimized by extending our model. Some of these questions may require a model with multiple classes of CFRs, which is an interesting direction for future work.

CHAPTER 3

PHASED ALERTING OF COMMUNITY FIRST RESPONDERS FOR CARDIAC ARREST

3.1 Introduction

3.1.1 Background

Survival for out-of-hospital cardiac arrest (OHCA) can be significantly improved through bystander efforts (Kragholm et al. 2017). To shorten the time to good-quality cardiopulmonary resuscitation (CPR), some emergency call centers use mobile phone technology to rapidly locate and alert nearby trained volunteers. A number of such community first responder (CFR) systems are active in, for example, the United States (Brooks et al. 2016), the United Kingdom (Smith et al. 2017) and Europe (Oving et al. 2019).

An alerting policy prescribes which volunteer to alert when. Simple alerting policies send all alerts at once. More complex alerting policies may use *phased* alerts, sending them in batches with time lags in between, to see if previous ones have been accepted. Such a policy is used by Hato Hone St John New Zealand, an ambulance provider that deploys the CFR system GoodSAM. Hato Hone St John New Zealand configured the system to dispatch CFRs in batches of 3 with time lags of 1 minute.

3.1.2 Importance

While for an individual patient, alerting many volunteers is beneficial, this leads to a high alert rate for volunteers as well as an increased likelihood of multiple volunteers arriving on scene, which may diminish their perceived contribution. Both aspects can lead to a long-term negative impact on volunteers' willingness to respond (volunteer fatigue), which would ultimately lower future OHCA survival. It is therefore important to find a good balance between future volunteer fatigue and current volunteer response times.

3.1.3 Goals of This Investigation

This chapter aims to assist CFR system managers in understanding the consequences of their chosen alerting policy. To that end, we define a large set of policies and four different KPIs (expected patient survival, coverage, number of alerts and number of redundant arriving volunteers). We evaluate these policies through Monte Carlo simulation using historical data from GoodSAM responses in New Zealand. Finally, we emphasize how the optimal choice depends on volunteer density. This is the first study to systematically evaluate different alerting policies.

3.2 Methods

3.2.1 Setting

The process of both CFR and EMS response is depicted in Figure 3.1. There, the name 'EMS response time' reflects the duration that is common for EMS providers to measure, and for which they often face targets and reporting obligations. For clarity, we have also added a 'GoodSAM response time', which starts at the moment the CFR system is activated.

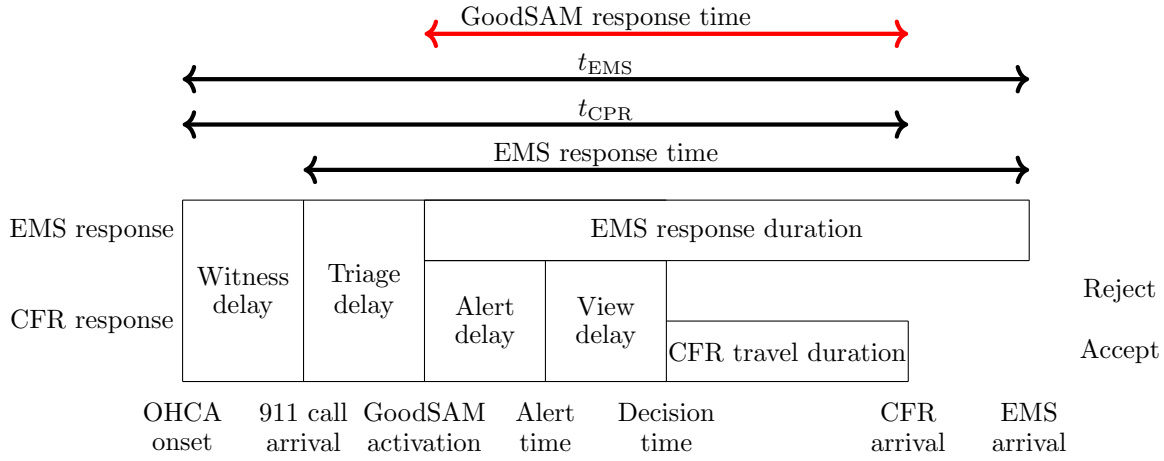


Figure 3.1: CFR response process.

EMS response times are measured from the moment of call arrival, so they are not ideal for estimating OHCA survival: the time that has passed since OHCA onset is more relevant. To that end, we define another duration, t_{EMS}: the duration between OHCA onset and EMS arrival. Similarly, we define t_{CPR}: the duration between OHCA onset and the arrival of the first responder, regardless of whether this is a volunteer or EMS. This is summarized in Figure 3.1.

3.2.2 Study Design

We defined a number of alerting policies, see Section 3.2.3. To estimate their performance, we used historical records of GoodSAM responses in New Zealand (Section 3.2.4) and applied them in a Monte Carlo simulation (Section 3.2.5) to quantify the GoodSAM response time (see Figure 3.1). To advance generalizability, we defined three different studies that correspond to different volunteer densities: low (10 volunteers within 1km), medium (30 volunteers within 1km) and high (100 volunteers within 1km). These densities roughly correspond to what would be found in Auckland, New Zealand if 0.2% of the inhabitants signed up (low), what is recommended (Stieglis et al. 2020) (medium) and our estimate for one of the most volunteer-dense regions worldwide (high). For each policy and each density, we use the GoodSAM response time obtained from simulation and convert them to expected values of four key performance indicators (KPIs, see Section 3.2.6).

3.2.3 Policies

We consider so-called alerting policies, which may base the decision to send an alert on the amount of time that has passed, and/or the responses that are received from previously alerted volunteers. Our policies only alert a more distant volunteer when all closer volunteers are already alerted. This helps to narrow down the set of all policies to realistic and promising ones. We further assume that no alerts are sent out after a volunteer accepts an alert, or 10 minutes have passed after GoodSAM activation, whichever comes first.

We evaluate the following policies:

1. *Send all at time 0*: alert all available volunteers within the dispatch radius immediately upon GoodSAM activation, which we call time 0.
2. *Send n at time 0*: Alert n volunteers immediately at time 0, and send no alerts after.
3. *Keep-n-active*: Alert n volunteers at time 0 and replace every incoming reject with an additional alert.
4. *NZ current policy*: Alert 3 volunteers at time 0 and 3 additional volunteers every minute until one volunteer accepts the alert, or 10 minutes have passed.

For *send n at time 0*, we consider the values of n between 1 and 15. For *keep-n-active* we consider values of n between 1 and 10. This gives a total of 27 policies.

3.2.4 Data

We used data obtained from the Hato Hone St John Ambulance Service, which operates the GoodSAM app in New Zealand. The data covers all OHCAs in the country between 1-12-2017 and 30-11-2020, along with their location and 911 call arrival time. The data also contains which CFRs were alerted through GoodSAM, at what time, and what their location was at the time of notification (*alert time* in Figure 3.1). Moreover, the data contains the time at which the volunteer reacted (*decision time*) by either accepting or rejecting the alert. Finally, the time volunteers arrived on scene (*CFR arrival*) is determined by the system, based on GPS signal. The data contains a total of 29,307 CFR alerts, which were either accepted (4009), accepted and dropped later (1199), rejected (7925), or not seen (16,174). From the 4009 accepting volunteers, 1776 eventually reached the patient.

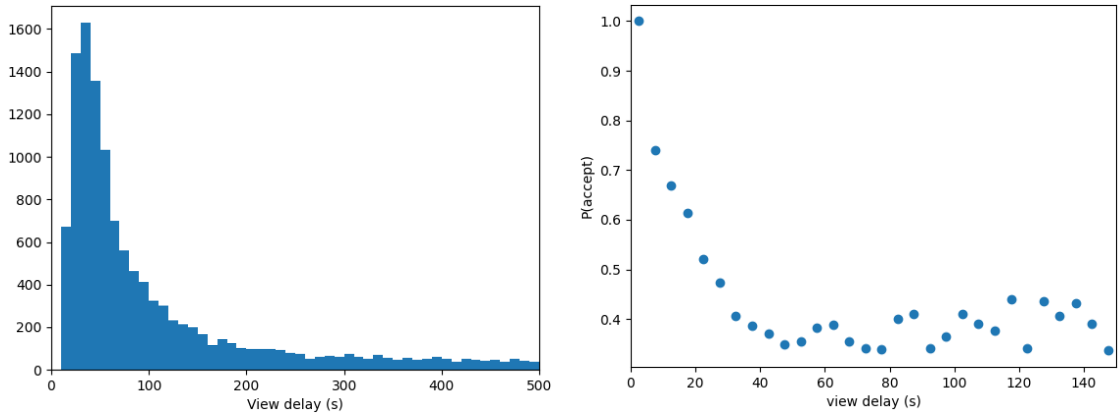


Figure 3.2: Historical view delays and acceptance probabilities as observed on GoodSAM in New Zealand.

For the view delay, we measured the duration between the alert time and the decision time. We ignored entries for which the alert time was equal to 'nan' (thereby excluding 1 entry). If the decision time was not listed, which is the case for not-seen alerts, we set the view delay to infinity. We treated both accepted and accepted-but-dropped-later entries as accepts, and we treated rejected and not-seen as rejects. This led to an empirical acceptance rate of $\frac{4009+1199}{29{,}307} \approx 17.77\%$. Figure 3.2 shows the empirical view delay distribution, as well as the estimated empirical probability of an accept, given a certain view delay. Not-seen alerts are not visible in this plot as these have an infinite view delay.

The CFR's travel duration depends on their distance from the patient. Locations and travel times were only recorded for *alerted* CFRs, so we cannot observe information on non-alerted individuals. However, in our counterfactual simulations those CFRs *may* be alerted, so we estimate their distances using Monte Carlo simulation. For each CFR that is within 1km of the patient, we simulate their location uniformly at random in this circle. This is also at the heart of a so-called Poisson process, which was previously proposed as a suitable model for CFR locations (van den Berg et al. 2024a) and used to represent the randomness

of OHCA locations over time (Skogvoll and Lindqvist 1999).

To translate distance into travel time, we used empirical data on volunteer travel speed, conditioning on the distance between volunteer and patient. We excluded 14 observations with a travel time of 0 seconds. We discretized the distances in 100-meter blocks and for each block, took the median speed of all volunteer responses in the data. We excluded the data for responses between 0 and 100 meters as we considered these less reliable. We then performed linear regression on the remaining medians to estimate the relationship between distance and speed.

3.2.5 Simulation

Our Monte Carlo simulation considers two sources of uncertainty: 1) the distance between the volunteers and the patient, and 2) the view delay and reply of alerted volunteers. The simulation evaluates 1,000 different OHCAs. For each OHCA, their CFR view delays and corresponding replies were bootstrapped 10,000 times from the empirical data. Each view delay and reply were jointly sampled to maintain the statistical dependence between the time a volunteer takes to respond and the likelihood that the reply is 'accept' (see Figure 3.2). This calculation estimates all KPIs with very narrow confidence intervals.

3.2.6 Outcomes

We consider two KPIs related to the quality of care and two KPIs related to the inconvenience caused to volunteers.

Coverage: fraction of incidents that receive a CFR response within a given time threshold. We define this on the 'GoodSAM response time' (see Figure 3.1) and use a threshold of 5 minutes.

Lives saved: the number of OHCA patients that are expected to survive until hospital discharge, in New Zealand, annually. This number is estimated using the relationship between time to CPR, time to EMS (t_{CPR} and t_{EMS} in Figure 3.1) and the probability of survival as modeled by Waalewijn et al. (2001). We assumed a t_{EMS} of 13 minutes as this is Hato Hone St John New Zealand's target for urgent requests (Ministry of Health 2020). To estimate t_{CPR}, we further need to estimate the 'Witness delay', which we assumed to be 1 minute, and the 'Triage delay', for which we used 124 sec (the median of the data).

The resulting survival percentage was converted to lives saved, by multiplying it by 5,141, which is the nationally observed number of OHCAs per year (Bray et al. 2022).

Number of alerts: the average number of alerts sent per incident.

Redundant arrivals: the average number of volunteers arriving on scene after the first responder, per incident.

3.3 Results

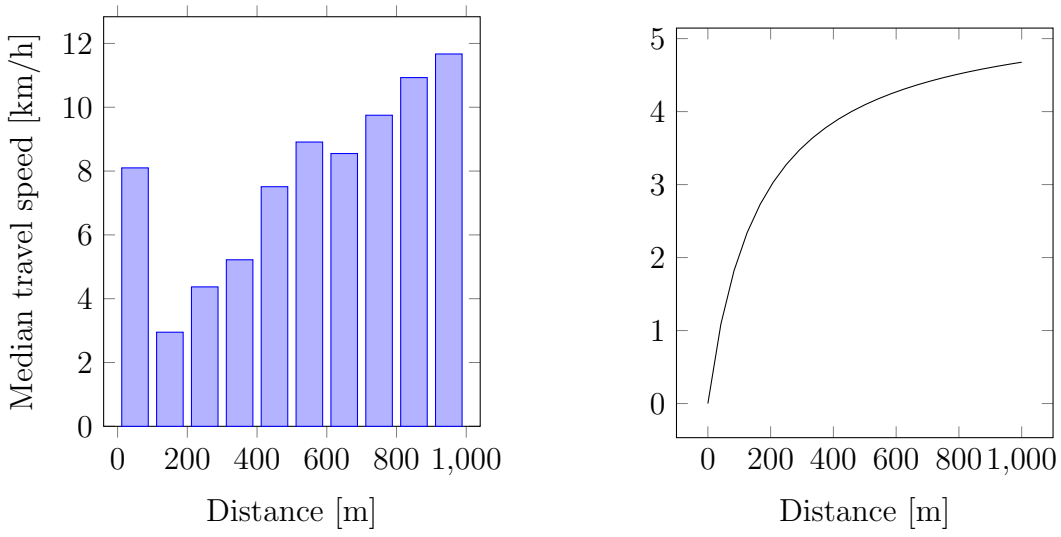
3.3.1 Parameter Estimation

Figure 3.3a shows the observed median travel speeds for different distances, which increase with distance. Linear regression of travel speeds in km/h, y, on distances in meters, x, based only on the distances above 100 meters has an R^2 of 0.96 and yields the relationship

$$y = 1.83 + 0.0108x. \tag{3.1}$$

The resulting relationship between the distance in meters and travel time in minutes, plotted in Figure 3.3b, is

$$travelTime(x) = \frac{60x}{1830 + 10.8x}. \tag{3.2}$$



(a) Median travel speed depending on distance to patient.

(b) Estimated travel time in minutes depending on distance to patient.

Figure 3.3: Distance-dependent travel times estimated based on historical data.

3.3.2 Performance of Policies

Tables 3.1-3.3 show the results for the different policies for a low, medium and high number of volunteers. For reference, alerting 0 volunteers would lead to 98 survivors, 0 redundant arrivals, 0 alerts and 0 coverage.

By only alerting the closest volunteer, the number of survivors can be increased to 120, 124, or 129, for 10, 30 or 100 volunteers within 1km, respectively. For coverage, the contribution of the closest volunteer is 11.7%, 13.4%, or 14.4%, respectively.

Comparing policies *Send n alerts at time 0*, with *Keep n alerts active* indicates that a similar level of survival can be obtained with 1 alert fewer at time 0, by replacing rejects with additional alerts. This reduces the impact of alerts on volunteers (both in terms of redundant arrivals as well as the number of alerts) without sacrificing survival.

The current policy used by GoodSAM in New Zealand on average leads to sending out 9.0 alerts per incident, if sufficiently many volunteers are within 1km from the patient. The corresponding number of expected survivors per year is 172, 190, or 212 for a low, medium and high number of volunteers. Coverage as a result of this policy is estimated at 26.6%, 35.3%, or 50.6%, respectively.

The policy that sends an alert to all volunteers at time 0 by definition yields the highest estimated survival (191, 231 and 259 survivors per year) and coverage (41.9%, 81.0% and 99.6%), but also gives the highest estimated number of redundant arrivals (0.918, 4.333 and 16.77) and alerts (10, 30, 100).

3.4 Discussion

The trade-offs between the KPIs in Tables 3.1-3.3 help a CFR system manager identify a desirable dispatch policy. This choice will likely depend on the volunteer density. Our reported number of volunteers within 1km of the patient has a direct relationship with the volunteer density. The three cases we considered correspond to a density of 3.18 vol/km^2, 9.55 vol/km^2, and 31.83 vol/km^2, respectively. For reference, the literature recommends a volunteer density of at least 10 vol/km^2 (Stieglis et al. 2020).

Because a different density might lead to a different preferred policy, the GoodSAM system offers the ability to configure different rules for urban and rural areas. This is done by uploading KML files that represent parts of the map. This feature implies that potential insights from this chapter can readily be implemented in practice.

Alerting the same number of volunteers leads to higher expected survival and better coverage as volunteer density increases. This effect is visible throughout Tables 3.1-3.3, and also holds in generality, assuming all other system parameters remain equal. The reason is that the n-th closest volunteer is expected to be nearer to the patient when volunteer density is higher.

Comparing policies *Send 6 alerts at time 0* and *Send 10 alerts at time 0* shows different survival numbers for different volunteer densities. In Table 3.1 the difference is 19 lives, while in Table 3.3 it is 26 lives. The reason is that in Table 3.1, the 6th-10th volunteers are farther from the patient than they are in Table 3.3.

Perhaps more surprising is that the policy *Send 6 alerts at time 0* gives approximately the same survival probability as *Keep 5 alerts active*, and that this is

the case in each of the three tables. We have no intuition for this, but conjecture that this result would be found for any volunteer density within the entire range of 3.18 – 31.83 vol/km^2.

Tables 3.1-3.3 show that Hato Hone St John's current alerting policy offers room for improvement. For example, with 10 volunteers in the circle (Table 3.1), this policy is expected to yield 172 survivors and sends, on average, 7 alerts per patient. In Table 3.1 we identify that the policy *Send 6 alerts at time 0* achieves the same level of survival with a higher coverage and a lower number of alerts and redundant arrivals. Moreover, Table 3.2 and 3.3 show that switching from Hato Hone St John's current strategy to *Send 7 alerts at time 0* is expected to yield a few additional survivors while simultaneously improving the other three KPIs. This insight might be a reason to fine-tune settings in the GoodSAM system.

To consider the generalizability of these results to other countries, we emphasize which parts of this research were New Zealand-specific: the view delay, travel speed and acceptance probability. Acceptance rates in New Zealand were found to be similar to those reported for other CFR smartphone apps (Haskins et al. 2021). CFR travel speeds and view delays are not readily available for other countries in the literature, but the framework designed for this study allows for easy testing on other parameters. Should readers be interested in whether the same conclusions hold for their country or region, the authors encourage them to reach out and offer to run simulations with different data.

3.4.1 Limitations

A limitation of this study is that the time on-scene is based on the geolocation of the phone, which may have some error due to gaps in cell tower coverage. It is unknown how often this situation occurs. The stop-gap to this is that responders can also manually generate an on-scene time by pressing a button in the app, but responders do not always do this.

Another limitation is that we focused on volunteers who directly go to the patient to provide CPR. However, in practice, certain CFR systems differentiate alerts, directing some volunteers to immediately attend to the patient and others to retrieve an Automated External Defibrillator (AED). Such a system raises new questions; answering those questions would require an extension of our methods.

We explored results for three different volunteer densities, but a CFR system will have varying volunteer densities throughout the day and the region. It might then be worthwhile to differentiate the dispatch policy within one system through times of the day and parts of the region. Accurately evaluating such a CFR system would require running simulations with the correct mix of densities and policies, thereby introducing complexity and likely diminishing insight. Therefore we decided not to pursue that approach here.

3.4.2 Future Research

Although it is plausible that a high number of alerts and/or redundant arrivals reduce future CFR engagement, there is, to the best of our knowledge, no study that quantifies the magnitude of this effect. We welcome such studies, as they

might aid in designing even better dispatch policies. The current knowledge gap led us to refrain from recommending a specific dispatch policy, instead presenting a list of options from which a CFR system manager can choose.

Future research might address more sophisticated policies. For example, consider a CFR as inactive when they have not responded for some time, or let the policy depend on the real-time locations of volunteers. It would also be interesting to discover the value of incorporating volunteer-specific knowledge, such as their historical acceptance rate, in the decision of which volunteer to dispatch. Finally, one might design a separate night-time policy that is trained on lower acceptance rates.

3.4.3 Conclusions

A CFR system's dispatch policy affects the trade-off between the number of alerts, redundant arrivals, and patient survival. In comparison to sending all alerts immediately, it is beneficial to send a reduced number of alerts immediately plus additional alerts upon receiving rejections. By doing this, a policy can be constructed to improve Hato Hone St John's current policy on all four KPIs. This chapter quantified KPIs for varying volunteer densities, and allows CFR system managers to choose a dispatch policy that suits their region. We show that the best choice depends on volunteer density and how one wishes to trade off KPIs. As volunteer densities may vary over the day and between different parts of the region, the choice may even be differentiated within one system. The GoodSAM system currently already allows for such user-specified geographical policy differentiation.

We hope that this chapter will stimulate discussion on phased dispatching of

CFRs and pave the way for incorporating methods from the field of operations research in designing improved dispatch policies.

Policy	Coverage	Survivors per year	Nr of alerts	Redundant arrivals
Send all at time 0.	0.419	191	10.0	0.918
Keep 1 alerts active.	0.123	123	1.3	0.000
Keep 2 alerts active.	0.207	140	2.6	0.042
Keep 3 alerts active.	0.268	152	3.8	0.115
Keep 4 alerts active.	0.314	163	4.9	0.211
Keep 5 alerts active.	0.348	171	6.0	0.326
Keep 6 alerts active.	0.373	178	7.1	0.454
Keep 7 alerts active.	0.391	183	8.0	0.587
Keep 8 alerts active.	0.404	187	8.8	0.718
Keep 9 alerts active.	0.413	190	9.5	0.834
Keep 10 alerts active.	0.419	191	10.0	0.918
Send 1 alerts at time 0.	0.117	120	1.0	0.000
Send 2 alerts at time 0.	0.202	135	2.0	0.032
Send 3 alerts at time 0.	0.265	147	3.0	0.089
Send 4 alerts at time 0.	0.312	157	4.0	0.168
Send 5 alerts at time 0.	0.347	165	5.0	0.265
Send 6 alerts at time 0.	0.372	172	6.0	0.375
Send 7 alerts at time 0.	0.391	178	7.0	0.498
Send 8 alerts at time 0.	0.404	183	8.0	0.630
Send 9 alerts at time 0.	0.413	187	9.0	0.771
Send 10 alerts at time 0.	0.419	191	10.0	0.918
NZ current strategy.	0.266	172	7.0	0.388

Table 3.1: Results for $n = 10$ volunteers within 1 km, based on 1000 simulated volunteer locations, with 10000 view delay and acceptance simulations each. The maximum confidence interval halfwidth as a fraction of the mean among all these numbers was 0.012 (ignoring those cases where the mean was zero). Policies that send more than 10 alerts at time 0 are omitted, as there are only 10 eligible volunteers in this situation.

Policy	Coverage	Survivors per year	Nr of alerts	Redundant arrivals
Send all at time 0.	0.810	231	30.0	4.333
Keep 1 alerts active.	0.149	128	1.3	0.000
Keep 2 alerts active.	0.259	148	2.6	0.041
Keep 3 alerts active.	0.345	164	3.8	0.114
Keep 4 alerts active.	0.414	176	4.9	0.211
Keep 5 alerts active.	0.471	186	6.0	0.326
Keep 6 alerts active.	0.518	193	7.1	0.456
Keep 7 alerts active.	0.559	200	8.2	0.596
Keep 8 alerts active.	0.593	205	9.2	0.743
Keep 9 alerts active.	0.623	209	10.2	0.898
Keep 10 alerts active.	0.649	213	11.2	1.057
Send 1 alerts at time 0.	0.134	124	1.0	0.000
Send 2 alerts at time 0.	0.241	143	2.0	0.032
Send 3 alerts at time 0.	0.328	157	3.0	0.089
Send 4 alerts at time 0.	0.399	169	4.0	0.168
Send 5 alerts at time 0.	0.459	179	5.0	0.264
Send 6 alerts at time 0.	0.510	187	6.0	0.375
Send 7 alerts at time 0.	0.552	193	7.0	0.498
Send 8 alerts at time 0.	0.589	199	8.0	0.630
Send 9 alerts at time 0.	0.620	204	9.0	0.771
Send 10 alerts at time 0.	0.646	208	10.0	0.919
Send 11 alerts at time 0.	0.669	211	11.0	1.071
Send 12 alerts at time 0.	0.689	214	12.0	1.228
Send 13 alerts at time 0.	0.706	217	13.0	1.389
Send 14 alerts at time 0.	0.721	219	14.0	1.553
Send 15 alerts at time 0.	0.734	221	15.0	1.719
NZ current strategy.	0.353	190	9.0	0.597

Table 3.2: Results for $n = 30$ volunteers within 1 km, based on 1000 simulated volunteer locations, with 10000 view delay and acceptance simulations each. The maximum confidence interval halfwidth as a fraction of the mean among all these numbers was 0.009 (ignoring those cases where the mean was zero.)

Policy	Coverage	Survivors per year	Nr of alerts	Redundant arrivals
Send all at time 0.	0.996	259	100.0	16.767
Keep 1 alerts active.	0.167	134	1.3	0.000
Keep 2 alerts active.	0.297	159	2.6	0.041
Keep 3 alerts active.	0.400	178	3.8	0.114
Keep 4 alerts active.	0.483	193	4.9	0.212
Keep 5 alerts active.	0.552	205	6.0	0.327
Keep 6 alerts active.	0.609	214	7.1	0.456
Keep 7 alerts active.	0.658	222	8.2	0.596
Keep 8 alerts active.	0.698	228	9.2	0.744
Keep 9 alerts active.	0.733	233	10.2	0.898
Keep 10 alerts active.	0.763	237	11.2	1.056
Send 1 alerts at time 0.	0.144	129	1.0	0.000
Send 2 alerts at time 0.	0.263	151	2.0	0.032
Send 3 alerts at time 0.	0.362	169	3.0	0.089
Send 4 alerts at time 0.	0.445	184	4.0	0.168
Send 5 alerts at time 0.	0.515	196	5.0	0.265
Send 6 alerts at time 0.	0.575	205	6.0	0.376
Send 7 alerts at time 0.	0.626	213	7.0	0.498
Send 8 alerts at time 0.	0.670	220	8.0	0.631
Send 9 alerts at time 0.	0.708	226	9.0	0.771
Send 10 alerts at time 0.	0.741	231	10.0	0.918
Send 11 alerts at time 0.	0.769	235	11.0	1.071
Send 12 alerts at time 0.	0.794	238	12.0	1.228
Send 13 alerts at time 0.	0.816	241	13.0	1.389
Send 14 alerts at time 0.	0.835	244	14.0	1.552
Send 15 alerts at time 0.	0.852	246	15.0	1.718
NZ current strategy.	0.506	212	9.0	0.603

Table 3.3: Results for $n = 100$ volunteers within 1 km, based on 1000 simulated volunteer locations, with 10000 view delay and acceptance simulations each. The maximum confidence interval halfwidth as a fraction of the mean among all these numbers was 0.006 (ignoring those cases where the mean was zero.)

CHAPTER 4

LEARNING TO DISPATCH VOLUNTEERS TO OUT-OF-HOSPITAL MEDICAL EMERGENCIES

4.1 Introduction

Certain medical emergencies, such as out-of-hospital cardiac arrests (OHCAs), require urgent medical attention, and rapid initiation of treatment can significantly increase survival rates. Many countries, including 19 out of 29 in Europe (Oving et al. 2019), have adopted a so-called community first responder (CFR) system. GoodSAM and Pulsepoint are two such applications introduced in multiple countries, including the UK, New Zealand, and the USA. CFRs are volunteers with some level of medical training who are willing to help during emergencies. CFR systems complement traditional ambulance services, ensuring timely response.

CFRs install an application on their smartphone that tracks, through GPS, their location and shares that location information with an application running in the emergency call center. When a medical emergency arises, the emergency call center can identify registered CFRs who are close to the medical emergency, and alert selected CFRs via the smartphone application. Besides CFRs, ambulances are dispatched as usual. While ambulances may be expected to follow instructions, alerted CFRs are not obligated to respond. However, those who choose to do so can proceed to the scene and provide first aid within the scope of their training until the independently dispatched ambulance crew arrives to take over. The central goal is to reduce morbidity and mortality by reducing response times. The CFR response process is summarized in Figure 4.1. There, the view delay is the time from the volunteer receiving the alert to responding, which can be indefinitely long

if the volunteer chooses to ignore the alert.

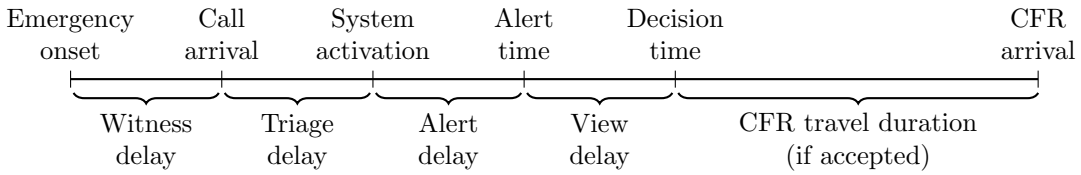


Figure 4.1: CFR response process.

Systems such as GoodSAM cope with inherent uncertainty regarding whether CFRs will accept their alert. From the perspective of a single patient, it is desirable to notify *all* nearby CFRs to minimize the arrival time of the first-arriving responder. However, this strategy may lead to volunteers receiving excessive alerts or an overwhelming number of volunteers arriving at the scene, both of which are believed to decrease the probability of volunteers accepting future alerts. This issue, known as volunteer fatigue, is a common challenge encountered by various volunteer schemes (Manshadi and Rodilitz 2020). Thus, it is important to balance a trade-off between the current CFR response time and the future CFR acceptance rate.

A dispatch strategy adopted by some CFR systems is to send *phased alerts*. Under this strategy, alerts are sent in batches with time lags in between, and once an acceptance is received from a CFR, future batches are not sent. The party managing GoodSAM in New Zealand (ambulance provider Hato Hone St John New Zealand) uses batches of three alerts and a 60s time lag between batches. While reasonable, this strategy does not explicitly consider volunteer fatigue. Moreover, the batch sizes and time lags are fixed and do not take into account the real-time locations of nearby CFRs, which are known at the time of the incident. To the best of our knowledge, to date, there are no CFR systems that use situation-dependent dispatch rules

In contrast to current practice, we present methods for determining which volunteers to alert and when for each incident, effectively balancing response times and volunteer fatigue and using real-time information on CFRs' locations and dispatched ambulances. We first introduce a Markov Decision Process (MDP) formulation aimed at optimizing response time and/or survival rates while penalizing volunteer fatigue. We then analyze the characteristics of MDP optimal policies. Recognizing the computational demands of obtaining a MDP optimal solution with increasing problem size, we offer an alternative data-driven approach using simulation and classification algorithms. This method can quickly identify the "best" policy among a set of pre-defined policies for each incident, facilitating real-time decision-making and easy deployment in practice. Additionally, the selection process of this approach is highly interpretable compared to the MDP optimal solution. These two proposed methods complement each other well, as we can observe important characteristics from the MDP optimal solution and use those to construct our set of predefined policies for the selection strategy. Finally, we validate the performance of our proposed methods through a case study in Auckland, New Zealand.

The remainder of this chapter is organized as follows. Section 4.2 provides a review of the related literature. Section 4.3 elaborates on the volunteer dispatch problem and introduces the developed MDP formulation. Section 4.4 presents a data-driven approach, based on simulating data and training a classification tree. In Sections 4.5 and 4.6, the application of our methods is demonstrated within the context of Auckland, New Zealand. We conclude with a discussion on the managerial implications of our findings and outline potential directions for future research in Section 4.7. Appendix B provides supporting details.

4.2 Literature

This section reviews the operations research literature on volunteer dispatch. Section 4.2.1 summarizes the literature specific to CFR systems for OHCA patients. Section 4.2.2 reviews studies on volunteer dispatch in other domains. Finally, Section 4.2.3 reviews work that inspires our proposed tree-based volunteer dispatch strategy.

4.2.1 Community First Response for OHCA

In recent years, we observe an increasing number of papers focused on mathematical modeling of CFR systems. A common aim is to quantify the effectiveness of such systems, and methods include simulation (Khalemsky and Schwartz 2017, Lancaster and Herrmann 2021, Buter et al. 2024) and convex optimization (van den Berg et al. 2024a). These studies typically assume a predetermined dispatch strategy and vary another aspect, such as volunteer density or acceptance ratio.

In contrast to the papers above, there are also studies that investigate how the dispatch process should be designed. In this context, a static approach entails optimizing a fixed dispatch radius (Slaa 2020). Another approach is to focus on the time-dynamic aspect of the problem; alerts could be sent sequentially. In Henderson et al. (2022) we ask how this could be optimized. One way of studying the sequential process is through Monte Carlo simulation, which allows the user to compare a handful of policies. This is done in van den Berg et al. (2024b), which evaluates three metrics: expected survival, number of alerts sent, and redundant arrivals. Finally, an approach not yet explored in the literature is applying mathematical optimization to design a dynamic alert policy.

Some CFR systems include task differentiation: they prescribe which volunteer should pick up a public access defibrillator and who should travel directly to the patient. Matinrad et al. (2019) attempts to optimize such task assignment. Matinrad et al. (2021b) and Matinrad and Granberg (2023) expand on that effort with uncertain task compliance, where volunteers sometimes divert from their assigned task, either skipping or adding an unplanned defibrillator pickup to their journey. Incorporating public access defibrillator pickups requires an extensive registry of their locations including up-to-date maintenance records - something many countries cannot offer. In particular, the CFR system that we have been working with does not include defibrillator pickups, and hence, for this chapter, task differentiation is out of scope.

4.2.2 Volunteer Dispatch in Other Domains

Volunteer fatigue has been explored in many different volunteer schemes. Sampson (2006) addressed the optimization of general volunteer labor assignments and showed that volunteer satisfaction is important for ensuring future participation. Manshadi and Rodilitz (2020) were inspired by a platform that recovers food from local businesses and donates it to nonprofit agencies, crowdsourcing the transportation to volunteers. Similar to CFR systems, this platform alerts multiple volunteers in the hope that at least one will respond positively. They expect excessive notifications to reduce volunteer engagement and identify a trade-off between notifying more volunteers for the current task and preserving them for future ones. In contrast to our work, which focuses on dynamic alert policies considering real-time volunteer response, they formulated the corresponding online volunteer notification problem as a generalization of online stochastic bipartite matching, with the

objective of minimizing the number of missed tasks. To capture a volunteer's adverse reaction to excessive notifications, they assume that a notification triggers a random period of inactivity.

Other studies also explore volunteer dispatch policies through optimization. Ata et al. (2019b) discuss volunteer gleaners who harvest leftover crops donated by farmers. Their paper aims to increase operational effectiveness by better managing uncertain food and labor supplies. It optimizes a dynamic volunteer-staffing policy, specifying staffing levels for each crop type based on the number of available gleaners and the backlog of donations. Escallon-Barrios et al. (2024) investigate the scheduling of a mixed workforce consisting of centrally scheduled paid staff and self-scheduled volunteers. By accounting for volunteers' responses to changes in scheduling options, they reduce over-covered and under-covered time slots while keeping volunteers engaged

Similar to CFR systems, spatial crowdsourcing involves tasks with assigned spatial locations that workers or volunteers must travel to (Kazemi and Shahabi 2012). In contrast to our work, which focuses on minimizing the time to complete a single task, most spatial crowdsourcing literature addresses a large number of tasks with goals such as task matching, quality control, incentive mechanism design, or privacy protection, as surveyed in Tong et al.. One major difference that separates CFR systems from the aforementioned works is the extreme time sensitivity of volunteer dispatch for OHCA, where tasks must be completed within a very short period. Additionally, due to the infrequent occurrence of OHCA incidents, the sparsity of incidents in both time and space prevents us from formulating our dispatch problem as an online matching problem, leading us to focus on single-incident scenarios.

4.2.3 Machine Learning and Policy Selection for Real-time Decision Making

In recent years, there has been significant work demonstrating the potential of using machine learning to support real-time decision-making. van der Hagen et al. (2022) explored the use of machine learning (ML) to support time slot management decisions in online grocery delivery. Their primary focus is on dynamically checking the feasibility of inserting a new customer order into existing delivery schedules. They transformed the feasibility check from solving a complex vehicle routing problem with time windows into a binary classification problem.

There are also studies focused on adaptively selecting the best policy among a finite set of alternatives. Bertsimas and Dunn considered the use of classification trees in this context due to their high interpretability and introduced the Optimal Classification Tree (OCT) method. This method formulates the decision tree creation problem as a mixed-integer optimization problem to create policy trees. A related strand of work is contextual ranking and selection (Alban et al. 2021, Cakmak et al. 2021, Keslin et al. 2022, Du et al. 2023), which determines the best policy from a predefined set based on simulation results. Our work incorporates both approaches by using simulation to evaluate each policy among the set of alternatives, then formulating the real-time decision as a classification problem using decision trees.

4.3 Markov Decision Process Formulation

In this section, we present a finite-horizon MDP formulation of the volunteer dispatch problem in discrete time with fixed time increments. Let $\mathcal{N} = \{1, \ldots, N\}$ denote the set of volunteers within the dispatch radius at the time of volunteer system activation, ordered by their travel time, from shortest to longest. We assume a single ambulance is dispatched and that the travel times for both the volunteers within the dispatch radius and the dispatched ambulance are known and deterministic at the time of volunteer system activation. Let T be the number of time steps it takes for the dispatched ambulance to reach the patient, and D_n be number of time steps for the n-th closest volunteer. Without loss of generality, we consider only volunteers with $D_n \leq T$ since there is no point in dispatching volunteers that, if dispatched immediately, still arrive after the ambulance. The volunteer system activates at time step $t = 0$, and the planning horizon ends when the ambulance reaches the patient at $t = T$.

4.3.1 State Space

The state space consists of three parts: t, $\mathbf{v_t} = \{v_t^1, \ldots, v_t^N\}$ and r_t. Here, t represents the current time step. Each $v_t^n, n \in \mathcal{N}$ provides the status of each volunteer at t. The term r_t indicates the number of time steps remaining for the fastest responder to reach the patient, whether the fastest responder is a volunteer or the dispatched ambulance.

The status of each volunteer, v_t^n, can be one of four possible states: 0, 1, 2, or 3, where 0 indicates that the volunteer has not been alerted; 1 indicates that the volunteer has been alerted but has not yet responded; 2 indicates that the

volunteer has been alerted and has accepted the alert; and 3 indicates that the volunteer has been alerted and has rejected the alert.

At the time of activation of the volunteer system, it is assumed that no volunteers have been alerted, so $r_0 = T$, indicating that the dispatched ambulance is the fastest responder at the moment. States with $r_t = 0$ are considered terminal states because they indicate that the first responder has already arrived at the scene, and no further actions are needed.

4.3.2 Action Space

Under the assumption of homogeneous volunteers, we focus on policies that prioritize alerting the nearest volunteer first. A simple interchange argument shows that this approach is optimal, because alerting the closer volunteer increases the chances of an early arrival and, thus, the survival rate. We further constrained ourselves to policies that permit multiple alerts to be sent out only at the volunteer system activation ($t = 0$), with at most one alert sent at each subsequent time step. At the time of system activation, the action space is thus given by $A_0 = \{0, 1, \ldots, N\}$, indicating how many volunteers to alert initially. For each subsequent time step, the action space is $A_t = \{0, 1\}$, where 0 indicates that no additional alerts are issued, and 1 indicates that the next closest unalerted volunteer should be alerted.

4.3.3 Transition

Let s_t be the state at time t. We denote $p(s_{t+1}|s_t, a_t)$ as the probability of transitioning from s_t to s_{t+1} following action a_t. The evolution of state s_t follows the

action a_t and is influenced by a random element $\omega(s_t, a_t)$ that characterizes volunteer responses. If $a_t = 0$, the status of all unalerted volunteers remains unchanged. If $a_t = 1$ and the next closest unalerted volunteer is n, then v^n_{t+1} transitions from 0 to 1, indicating that the volunteer has been alerted. Each alerted volunteer who is yet to respond has a fixed probability p_a of accepting the alert and p_b of rejecting it during each time step, with $1 - p_a - p_b$ being the probability of no response from the volunteer during the time step. Once a volunteer accepts or rejects an alert, their status does not change further.

4.3.4 Reward Function

We maximize the overall effectiveness of volunteer dispatch, which is modeled as the survival probability of the patient subtracting the volunteer fatigue. Volunteer fatigue is proportional to the number of redundant volunteers (those arriving after the first responder has already reached the scene). The system only incurs a reward upon entering terminal states, i.e., when the first responder has reached the patient. To define those rewards, denote $f(s)$ as the expected survival probability and $g(s)$ as the expected number of redundant volunteers upon entering terminal state s. We quantify the patient survival following Waalewijn et al. (2001), which was fit on data where both ambulances and volunteers responded, and volunteers did not carry defibrillators. This function takes the time to CPR (the time of first response, whether by volunteer or ambulance) and the time to EMS (time of arrival of the ambulance) from patient collapse in minutes as inputs, and is given by

$$\phi(t_{CPR}, t_{EMS}) = (1 + \exp\{0.04 + 0.03 t_{CPR} + 0.14(t_{EMS} - t_{CPR})\})^{-1}. \quad (4.1)$$

Let γ represent the penalization factor for the number of redundant volunteers,

and let $h(s_t, a_t, s_{t+1})$ denote the reward for transitioning from state s_t to s_{t+1} following action a_t. Then

$$h(s_t, a_t, s_{t+1}) = \begin{cases} f(s_{t+1}) - \gamma g(s_{t+1}) & r_{t+1} = 0 \\ 0 & \text{otherwise} \end{cases}. \tag{4.2}$$

We assume that all outstanding alerts without any responses are considered canceled 15 minutes after system activation. Consequently, only those volunteers who accept alerts within 15 minutes of system activation contribute to the count of redundant volunteers. Volunteers who arrive simultaneously with the ambulance are also counted as redundant.

4.3.5 Optimal Dispatch Policy

Let τ denote the random time representing the arrival of the first responder, defined as $\tau = \min\{t \geq 0 \colon r_t = 0\}$. For all $t = 0, 1, \ldots, T-1$ and states $s_t \in \mathcal{S}$, the value function following a policy π is given by

$$V^{\pi}(s_t) = \mathbb{E}_{\pi}\left[\sum_{t=t}^{\tau-1} h(s_t, \pi(s_t), s_{t+1}) \big| s_t\right] \tag{4.3}$$

$$= \sum_{s_{t+1}} p(s_{t+1}|s_t, \pi(s_t)) \left[h(s_t, \pi(s_t), s_{t+1}) + V^{\pi}(s_{t+1})\right]. \tag{4.4}$$

We determine the optimal volunteer dispatch strategy π^* following the Bellman optimality equation

$$\pi^*(s_t) = \text{argmax}_{a \in A_t} \left\{ \sum_{s_{t+1}} p(s_{t+1}|s_t, a) \left[h(s_t, a, s_{t+1}) + V^{\pi^*}(s_{t+1})\right] \right\}. \tag{4.5}$$

With this setup, we can establish the following properties of the optimal value function and policy. Proofs are provided in the B.1.

A Pareto-optimal frontier provides a set of solution points that represent the best trade-off between multiple objectives. We consider it because our objective involves a trade-off between patient survival and the number of redundant volunteers. Proposition 4.3.1 demonstrates that the Pareto-optimal frontier in the volunteer dispatch problem is concave, indicating that each additional redundant volunteer on-scene yields a progressively smaller increase in patient survival.

Proposition 4.3.1. *The Pareto-optimal frontier is concave with respect to the expected volunteer fatigue for any fixed OHCA.*

Propositions 4.3.2 and 4.3.3 establish key characteristics of the optimal dispatch policy.

Proposition 4.3.2. *Once someone accepts, the optimal policy π^* will not send out any additional alerts.*

Proposition 4.3.3. *If all alerted volunteers have rejected alerts by time t and, under the optimal policy π^*, no new alert is sent at time t, then no alerts will be sent out afterward.*

While the policies obtained from the above MDP model are optimal, their computational demands make real-time application infeasible, especially for time-sensitive medical emergencies like out-of-hospital cardiac arrests. Additionally, these policies can be difficult to interpret. Therefore, in the next section, we introduce an alternative data-driven method that is both fast and interpretable.

4.4 Selection Strategy Using a Classification Tree

In this section, we present an alternative data-driven approach that utilizes both simulation and classification algorithms. First, we introduce a finite set of predefined policies, each of which is considered to be effective, at least for some incidents. Next, we describe a simulation model used to evaluate the effectiveness of each policy. Finally, we outline a selection strategy using classification algorithms to quickly identify the "best" policy from the predefined set for each incident, facilitating real-time decision-making and practical deployment. This approach offers greater interpretability, making the policy selection process more transparent and understandable.

4.4.1 Dispatch Strategies and Pre-Defined Policies

Here, we distinguish between *stratgies* and *policies*. A *policy* defines how CFRs are dispatched for a single incident, while a *strategy* determines which policy to use for each incident. CFR systems generally apply a *static* strategy, meaning they use a fixed policy for all incidents, even though the policy can adapt based on real-time CFR responses. In contrast, we propose an *adaptive* strategy, which selects the appropriate policy at the time of dispatch based on real-time information, such as CFR locations, assigned ambulances, and volunteer system triage delays.

Our approach converts the problem into a multi-class classification problem. To achieve this, we first define a finite set of policies that will serve as the classes. Although computing the optimal policy from the MDP model is computationally intensive, we have identified key characteristics from the MDP optimal policy that we want to incorporate into our predefined policies.

A key characteristic is that no further alerts should be sent once a volunteer has accepted an alert. Additionally, if a rejection is received, the next alert should be sent immediately rather than waiting, as suggested by Propositions 4.3.2 and 4.3.3. We also observed that it is often unnecessary to alert additional volunteers if there has been no response. After discussions with people who currently operate CFR systems, we came up with the following alert policies, starting the time count from the moment of CFR system activation:

1. *Send all at time 0*: alert all available volunteers within the dispatch radius immediately .
2. *Send n at time 0*: Alert n volunteers immediately at time 0, and send no alerts thereafter.
3. *Keep n active until time t*: Alert n volunteers at time 0 and replace every incoming reject with an additional alert. After time t, stop sending alerts.

We recognize the fact that many different types of dispatch policies exist. Our approach is flexible and can accommodate various dispatch policies. However, for the purposes of this chapter, we focus on this predefined set. These policies are intuitive, readily implemented and incorporate key characteristics observed from the DP optimal policy. In addition, the Keep n active until time t policies resemble the *phased alert* policies that are commonly used in CFR systems. We believe these policies are sufficiently robust and, when selected in an incident-specific manner, cover a reasonable range of strategies.

4.4.2 Simulation Model

We evaluate the performance of each dispatch policy using a discrete-event simulation model that includes the entire CFR response process:

1. CFR system activation.
2. Alerted CFR response: accepting, rejecting or ignoring the alert.
3. Issuing additional alerts if needed, following the specified policy.
4. CFR arrival on-scene to treat the patient if they arrive before the ambulance.
5. Ambulance arrival.

For each instance, we assume the dispatcher knows the estimated travel times of the assigned ambulance and the CFRs within the dispatch radius and that a CFR system triage delay has been realized.

We assume the CFR responses (including their answer and corresponding view delay) follow a known distribution $\mathcal{F}$. We used an empirical response distribution based on the data provided by Hato Hone St John Ambulance Service. More details on the dataset are included in Section 4.5.2.

If no CFRs are available, only the ambulance will respond. If multiple CFRs are available, we assume only the first responding CFR will perform the treatment. All CFRs arriving after the first responder (whether a CFR or the assigned ambulance) are considered redundant. Any outstanding alerts that receive no response will be considered canceled 15 minutes after system activation, as they will no longer be helpful for the patient.

4.4.3 Machine Learning Pipeline

Given the set of predefined policies, we aim to develop an *adaptive* strategy that can identify the "best" policy in real-time whenever an OHCA occurs. While any machine learning classification algorithm could be used for this purpose, we use a classification tree, partly inspired by Bertsimas and Dunn, to learn an interpretable, tree-based CFR dispatch strategy.

We define the CFR response time as the sum of the triage delay and the CFR travel time. The predictors we use for the classification tree include the number of CFRs within the dispatch radius, their corresponding response times, and the realized CFR system triage delay. We use CFR response times instead of CFR travel times as features because considering triage delays and travel times separately can be misleading and may result in unnecessary misclassification. For example, a large triage delay combined with a fast CFR travel time can still result in an overall fast response time. Therefore, it is more important to focus on the overall response time.

When learning this selection strategy, we assume a fixed ambulance travel time across all instances in the training set for simplicity and more intuitive results. However, our approach can be extended to include real-time estimates of ambulance travel time as an additional feature. We generated 10,000 instances as our training dataset, labeling the optimal dispatch policy for each instance using the simulation model described above. Additionally, we generated 2,000 instances as our testing set. More details on the instance generation process are included in Section 4.5.3, and we present the numerical results in Section 4.6.

4.5 Case Description

Auckland is the largest city in New Zealand, with a population of 1.4 million within the greater Auckland region. Our study focuses on the most urban areas with a population density exceeding 1000 people per square kilometer. The focus area, with 1.1 million residents, is divided into 287 so-called area units. These area units are the second smallest unit at which Statistics New Zealand collects data (Stats New Zealand Geographic Data Service 2016).

4.5.1 CFR Response Process

We follow the CFR response process as shown in Figure 4.1. The time from call initiation to volunteer system activation is defined as the *triage delay*. The time from system activation to the issuing of the first set of alerts is defined to be the *alert delay* and is assumed to be zero. The time from the volunteer receiving the alert to responding is called the *view delay*, which can be indefinitely long.

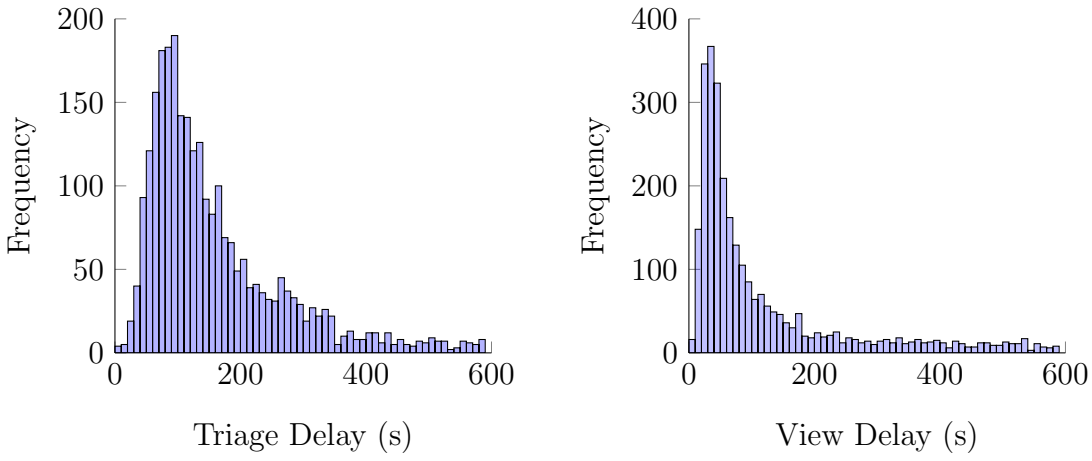
We assume that accepting CFRs respond on foot, and use the most conservative estimate of Slaa (2020) in the Netherlands to model their speed at 6 km/h, traveling as-the-crow-flies distances. We assume that CFRs are dispatched only if they are within 1 km of the patient and can reach the patient before the assigned ambulance, aligning with current practice. Consequently, the travel time for CFRs, assuming at least one is available within 1 km, is 10 minutes or less.

4.5.2 Empirical Data

We used data provided by Hato Hone St John Ambulance Service (SJAS), which operates the GoodSAM app in New Zealand. The SJAS has provided us with all OHCAs from 2013 to 2020, with their locations and 911 call arrival times. The dataset also includes details of CFR alerts issued from inception of the program in 2017 to 2020, with the time and location of each alert, the CFRs' responses (acceptance, rejection, or no response), and the arrival time at the scene based on GPS data.

Among the total 4157 OHCA incidents arising while the CFR system was operating and in the area of interest, 2827 had alerts sent to CFRs. Of the 6749 CFR alerts recorded, 698 were accepted, 263 were accepted but later dropped, 2026 were rejected, and 3762 received no response. If an alert was initially accepted but later dropped, it indicates that the CFR initially agreed to respond but later decided not to, possibly due to seeing someone else already assisting the patient or other reasons. We treat both outright acceptances and initial acceptances (dropped later) as acceptances, which leads to an empirical acceptance rate of 14%. We measured the triage delay from call arrival to the earliest alert time, and view delay from alert time to when CFRs make their decisions.

Figure 4.2 presents the empirical distributions of triage delays and view delays. We omitted 2 incidents with missing call arrival times for the histogram of triage delays. The median triage delay is 2 minutes. Among all responding CFRs, 95% responded within 10 minutes, and 84% responded within 5 minutes. The peak view delay is between 20s to 50s, with the peak of those who accepted alerts responding slightly faster than those who rejected. This suggests that CFRs who accept alerts tend to respond more quickly. If a CFR did not respond and ignored the alert,



(a) Closest CFRs

(b) Alerted CFRs

Figure 4.2: Left: Histogram of triage delays from the 2825 OHCA incidents in Auckland from December 2017 to December 2020. Right: Histogram of view delays from the 6749 alerts in Auckland from December 2017 to December 2020

the decision time (whether to accept or reject the alert) is not recorded in the dataset. In this case, we assign an infinite view delay, which will not be observed in Figure 4.2b.

4.5.3 OHCA Incidents Generation

In our Monte Carlo simulation, we sample the location of an OHCA using an area-unit specific incident rate that is estimated by integrating the SJAS dataset and demographics with socio-economic factors as discussed Dicker et al. (2019). One could instead use the empirical OHCA rate, but low counts in some area units have very high variability relative to their mean. More details and discussion on this estimation process can be found in Appendix A.7.

We assumed that 0.1% of the population in each area unit are registered CFRs, which allowed us to calculate the volunteer density per area unit. We obtain the

resident population in each area unit from the 2013 census (Stats New Zealand Geographic Data Service 2015a). We then applied the spatial Poisson point process model for CFR distribution as introduced in van den Berg et al. (2024a) to simulate the number and locations of volunteers within a 1 km dispatch radius.

For each alerted volunteer, we bootstrap their view delay from the SJAS dataset. To preserve any potential dependence between the view delay and volunteer response (acceptance or rejection), we sample these variables concurrently. In addition, we sample the triage delay for each incident from the SJAS data by excluding any instances where the triage delay exceeds 10 minutes, as volunteer response with a delay of this length is unlikely to contribute to patient survival.

4.5.4 Policies

We used the set of pre-defined policies discussed in Section 4.4. For *send n at time 0*, we consider the values of n between 0 and 10. For *Keep n active until time t* we consider values of n between 1 and 10 and set t equal to 3, which proved most fruitful in extensive preliminary experiments. This gives a total of 22 policies.

4.5.5 Performance Metrics

To measure the effectiveness of each volunteer dispatch policy, we measure the patient survival and subtract the volunteer fatigue. We quantify the patient survival using (4.1), which requires volunteer and EMS response time measured from patient collapse as input. In line with Waalewijn et al. (2001), we add a 1-minute delay from patient collapse to call initiation. We obtain the CFR response time

from call initiation from simulation. Additionally, we consider three different ambulance response times from call initiation: a long response time of 12 minutes, which SJAS aims to meet for 95% of incidents; a short response time of 6 minutes, targeted for 50% of incidents; and a median response time of 9 minutes. Volunteer fatigue is calculated based on the number of redundant CFR dispatches, applying a multiplicative penalization factor of 0.01.

4.6 Numerical Results

In this section, we present the performance of our proposed tree-based selection strategy through the case study described in Section 4.5. We first provide an example of the tree-based selection strategy in Section 4.6.1, then compare its performance with that of static strategies in Section 4.6.2. Recognizing that volunteer acceptance rates vary across different countries and regions, we examine the impact of the acceptance rate on dispatch policies in Section 4.6.3. Finally, we conclude the section with a comparison between the MDP optimal policy and the tree-based strategy in Section 4.6.4.

4.6.1 High Interpretability of Tree-Based Strategy

A key advantage of the proposed tree-based selection strategy over the MDP formulation is its high interpretability, which is especially valued in healthcare-related applications, including CFR systems.

Figure 4.3 shows an example of the tree-based selection strategy trained on generated instances in Auckland with a 12-minute EMS response time, where each

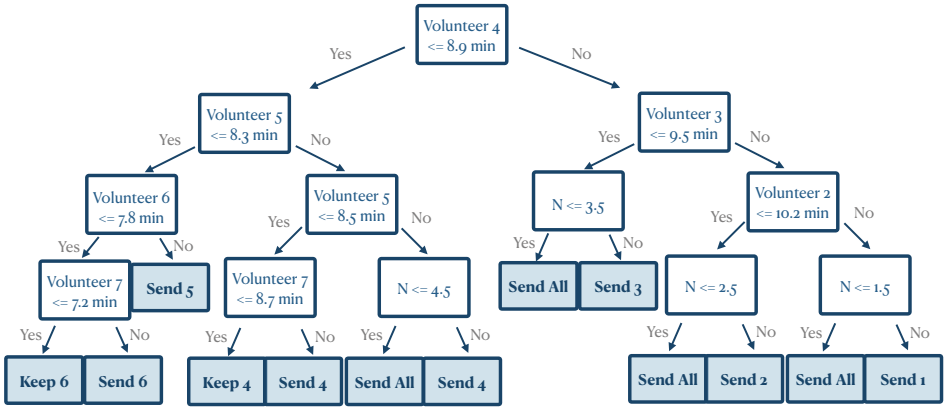


Figure 4.3: Tree-based selection strategy for instances in Auckland with 14% acceptance rate and 12-minute EMS response time measured from call arrival.

node represents a split in the feature space. The tree is easy to understand, as it selects policies that send out more alerts when there are more CFRs within the dispatch radius and closer to the patient. Conversely, on the right side of the tree, with a smaller number of CFRs or when the closest CFRs are far away from the patient, the strategy refrains from sending out additional alerts.

4.6.2 Adaptive Tree-Based Strategies vs. Static Strategies

We next compare the performance of our adaptive tree-based strategy with static strategies under three different EMS response times. We then focus on a 12-minute EMS response time, partly because SJAS aims to respond to 95% or more calls within 12 minutes. More examples of the tree-based strategies under different EMS response times are included in B.2.

Recall that our objective is defined as patient survival minus volunteer fatigue, with a penalization factor of 0.01. Table 4.1 provides the accuracy of our tree-based strategy on the testing set, measured as the fraction of test instances where

the tree-based strategy correctly identified the "best" policy among the predefined set. It also shows the average objective value of the tree-based strategy compared to the best static strategy under each scenario. The best static strategy is defined as the policy from our pre-defined set that yields the highest average performance across all test cases when applied uniformly.

EMS	Best Static Strategy	Tree Accuracy	Tree Obj	Best Static Obj
6 min	*Send 1 at time 0*	0.973	0.107	0.1062
9 min	*Send 3 at time 0*	0.933	0.0505	0.0497
12 min	*Send 4 at time 0*	0.812	0.0254	0.0249

Table 4.1: Accuracy and average objectives reported by the tree-based selection strategy versus the average objectives of the best static strategy among our pre-defined policies under a 14% volunteer acceptance rate.

We observed that as the EMS response time increases, the best static strategy issues more alerts. This is because shorter EMS response times provide less opportunity for CFRs to react, increasing the likelihood that CFRs will arrive after the ambulance and contribute to volunteer fatigue. As a result, shorter EMS response times often result in simpler dispatch policies being optimal, which leads to a simpler classification tree and higher accuracy in predicting the correct "optimal" strategy.

In all scenarios, we observed a 1-2% improvement in average performance across test cases by selecting an appropriate policy based on real-time information, even if the tree-based policy occasionally misclassifies the best policy for certain instances. Figure 4.4 illustrates the trade-off between patient survival and redundant volunteers for both our adaptive and static strategies under a 14% acceptance rate and a 12-minute EMS response time. The tree-based selection strategy lies outside the efficient frontier of the static strategy, because it can tailor the response to each OHCA, in contrast to static strategies.

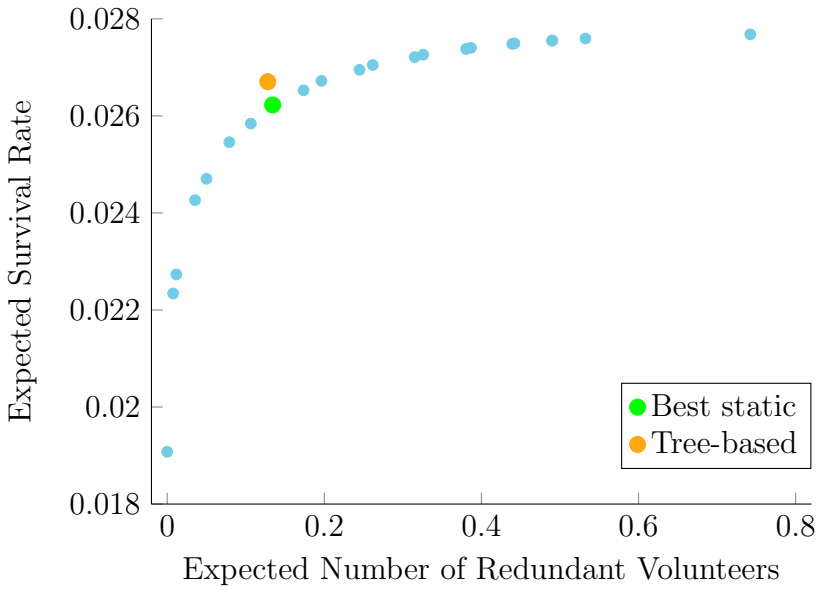


Figure 4.4: Trade-off between patient survival and number of redundant volunteers for static and adaptive tree-based strategies. Each blue dot represents the performance of one static strategy. The orange dot indicates the performance of the tree-based strategy. The green dot indicates the performance of the best static strategy, which is *Send 4 at time 0* under a 14% acceptance rate and a 12-minute EMS response time.

While the absolute difference in average performance over all test cases between the best static strategy and the selected tree-based strategy may be small, there can be significant improvements in individual instances. Figure 4.5 shows the percentage improvement in patient survival versus the percentage improvement in the objective of the tree-selected policy compared to the best static policy for each test case.

From Figure 4.5, the tree-based policy almost universally improves on the best static policy (almost all points lie to the right of 0 on the horizontal axis), but it can sometimes perform slightly worse than the best static policy, due to misclassification. In many instances, tree-selected policies improve the overall objective by sending out fewer alerts; these instances lie in the lower-right quadrant. In these instances a CFR response may not be very effective, and the savings from

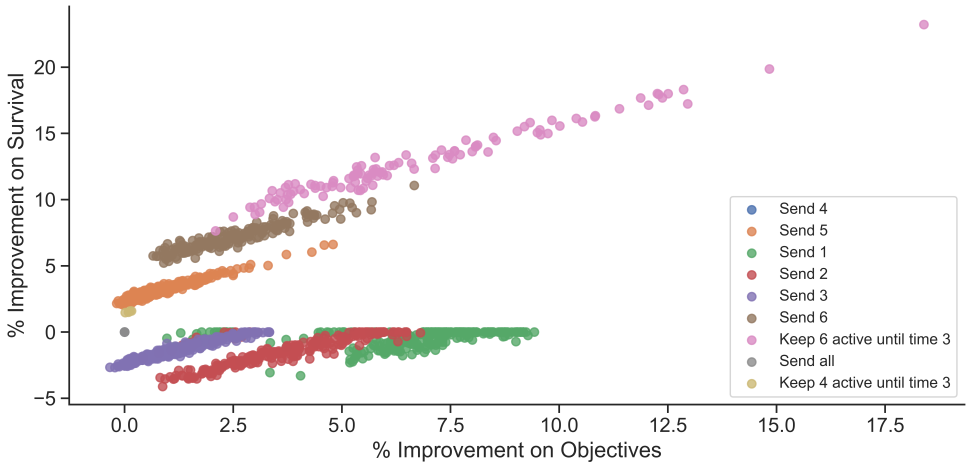


Figure 4.5: Percentage improvement in survival rate vs. the percentage improvement in the objective of the tree-selected policy compared to the best static policy for 2000 test cases under a 14% acceptance and a 12-minute EMS response time.

reducing redundant volunteers outweighs the decrease in patient survival. Additionally, the tree-based strategy is advantageous when there are some CFRs within the dispatch radius but even the closest CFR is still far from the patient. This is shown by the green points on the right end of Figure 4.5. In these instances, the tree-based strategy recognizes that a CFR response is unlikely to significantly improve patient survival, even with more alerts. When the potential improvement in survival is negligible, the strategy prioritizes conserving volunteer resources to reduce fatigue. Tree-based policies yield the most desirable results—improving both patient survival and overall performance—when there are a large number of CFRs within the dispatch radius who are relatively close to the patient. In these cases, the tree-based strategy selects a policy that sends out more alerts than the best static policy, leading to better outcomes that appear as points with positive improvements in survival in the plot.

4.6.3 Impact of Volunteer Acceptance Rate

In the previous section, we assumed a 14% acceptance rate by directly bootstrapping CFR responses from the data. However, in areas with long-established CFR systems, the acceptance rate is likely to be higher. Consequently, we explore the impact of a higher volunteer acceptance rate on optimal volunteer dispatch strategies and their performance.

We assume a 50% acceptance rate. When generating volunteer responses in our simulation, instead of directly bootstrapping from the empirical data, we first sampled a Bernoulli variable with a parameter of 0.5 to decide whether the response should be an acceptance or not. If it is an acceptance, we repeatedly sample the CFR response from the empirical data until we get an acceptance. This method boosts the acceptance rate from 14% to 50% while preserving the relationship between volunteer response and view delay.

Table 4.2 provides the accuracy of our tree-based strategy on the testing set and its average objective value, compared to the best static strategy under three different EMS response times and a 50% acceptance rate. Compared to Table 4.1, we can see that a higher acceptance rate leads to better performance for both the tree-based and best static strategies. The improvement over the best static strategy is also around 1-3% in all three cases.

EMS	Best Static Strategy	Tree Accuracy	Tree Obj	Best Static Obj
6 min	*Send 1 at time 0*	0.982	0.11	0.1086
9 min	*Send 2 at time 0*	0.913	0.0580	0.0566
12 min	*Keep 2 active until 3*	0.916	0.0330	0.0323

Table 4.2: Accuracy and average objectives reported by the tree-based selection strategy versus the average objectives of the best static strategy among our predefined policies under a 50% volunteer acceptance rate.

Under the same EMS response time as before, with a higher CFR acceptance rate, the best static strategy sends out fewer alerts. This is because a higher acceptance rate could potentially result in a greater number of redundant volunteers. Although not shown in the table, *Keep n active until t* policies are more likely to be optimal under a higher acceptance rate. With a higher acceptance rate, it is more beneficial to alert fewer CFRs initially to avoid volunteer fatigue. If the alerted CFR rejects, there is still a reasonable confidence that the next CFR will be available to ensure a fast response.

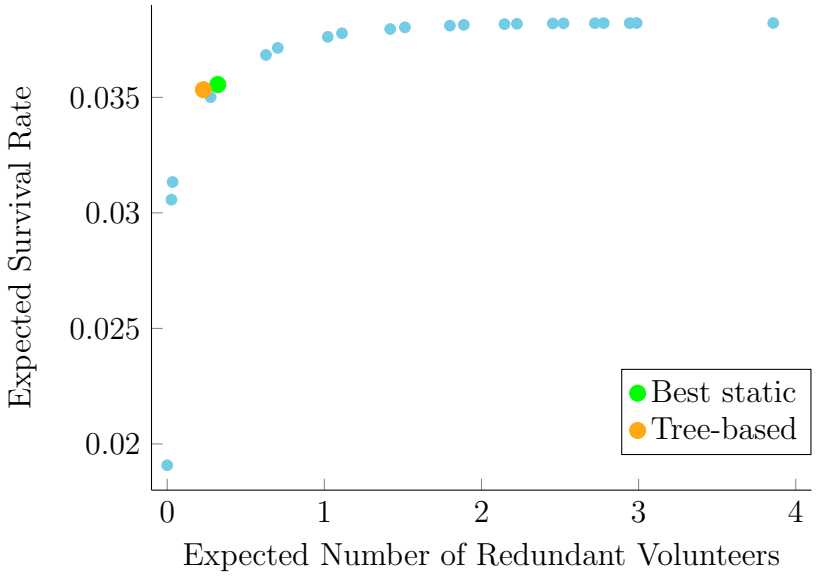


Figure 4.6: Trade-off between patient survival and number of redundant volunteers for static and adaptive tree-based strategies. Each blue dot represents the performance of one static strategy. The orange dot indicates the performance of the tree-based strategy. The green dot indicates the performance of the best static strategy, which is *Keep 2 active until 3* under a 50% acceptance rate and a 12-minute EMS response time.

Figure 4.6 illustrates the trade-off between patient survival and redundant volunteers for both our adaptive and static strategies under a 12-minute EMS response time and a 50% acceptance rate. Similar to Figure 4.4, the tree-based selection strategy lies outside the efficient frontier of the static strategies. The tree-based selection strategy decreases the redundancy rate relative to the best static policy,

but also slightly decreases patient survival. This is likely because, with such a high volunteer acceptance rate, the best static strategy already performs very well in terms of patient survival. As a result, the primary way to improve the overall objective is to reduce redundant volunteers without significantly compromising patient survival.

Figure 4.7 makes this observation more clear, showing the percentage improvement in patient survival versus the percentage improvement in the objective of the tree-selected policy compared to the best static policy for each test case under a 50% acceptance rate. The tree-based strategy improves upon the best static strategy in most cases by decreasing the number of redundant volunteers, at the cost of a decrease in patient survival. Only in a small number of cases, shown as the cluster of pink and grey points in the upper left corner of Figure 4.7, does the tree-selected policy improve both survival and the overall objective. This typically occurs when there is a high CFR density within the dispatch radius. Compared to Figure 4.5, Figure 4.7 more clearly demonstrates that the tree-based strategy is most advantageous in two scenarios: when there are a large number of CFRs who are relatively close to the patient (pink and grey points) and when there are some CFRs within the dispatch radius but the closest CFR is far from the patient (brown points).

4.6.4 MDP Optimal Policies vs. Tree-Selected Policies

Next, we show that with a good pre-defined set of policies to start with, the tree-based selection strategy can perform nearly "optimally." To be more consistent with the MDP setup, the tree-based selection strategy was trained on 10,000 instances with the same instance generation process as described in Section 4.5.3.

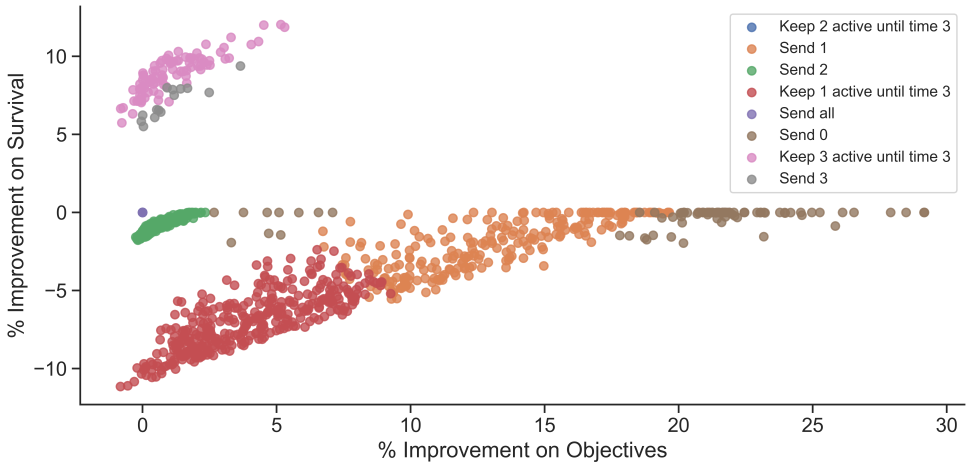


Figure 4.7: Percentage improvement in survival rate vs. the percentage improvement in the objective of the tree-selected policy compared to the best static policy for 2000 test cases under a 50% acceptance and a 12-minute EMS response time.

However, given the computational time of MDP optimal policies, we restrict ourselves to instances with 6 or fewer volunteers that can reach the patient before ambulance arrival. We assumed a volunteer acceptance probability of 0.00105 and a rejection probability of 0.00221 at each 5s time step so that the probability of accepting and rejecting within 15 minutes matches the empirical rates in the dataset.

Figure 4.8 compares the performance of MDP optimal policies versus tree-selected policies for 1000 test instances under a 12-minute EMS response time. In most instances, the tree-selected policy matches the performance of the MDP policy, which is optimal under this setting. The MDP optimal policy only significantly outperforms the tree-selected policy when misclassification occurs. The strong performance of our tree-based selection policy in this setting may be due to the fact that our pre-defined set of policies was constructed using characteristics observed from the MDP optimal policies. This suggests that with a well-chosen set of pre-defined policies that sufficiently covers the space of effective volunteer

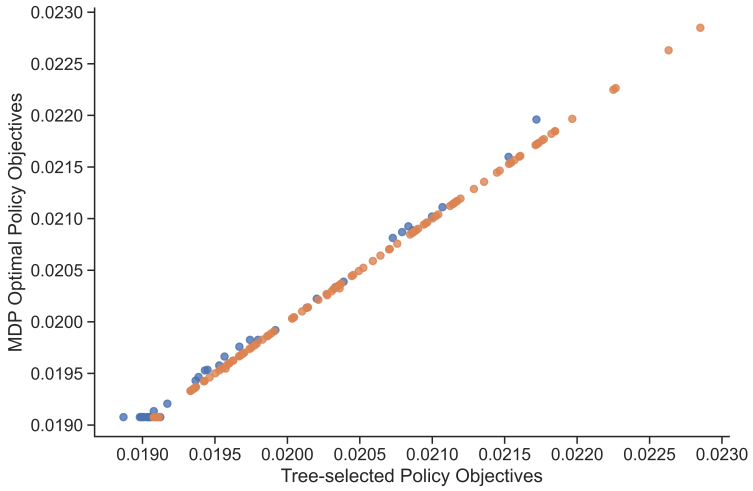


Figure 4.8: Performance of MDP optimal policies vs. performance of tree-selected policies on 1000 test cases under a 12-minute EMS response time and a 14% acceptance rate within 15 minutes of CFR system activation. The blue dots represent instances where the tree-based strategy misclassifies the "best" policy among the pre-defined set.

dispatching policies, the tree-based selection strategy can perform exceptionally well.

4.7 Conclusion and Discussion

In this chapter, we introduced a Monte Carlo simulation model for evaluating CFR dispatch policies. We presented an MDP formulation and a data-driven method using both simulation and a classification tree to learn the "optimal" CFR dispatch strategies. We demonstrated the performance of our proposed method through a comprehensive case study in Auckland in Section 4.5. While the results and insights for our tree-based strategy are specific to Auckland (Sections 4.6.1 and 4.6.2), the same selection strategy approach can be applied to other regions with data or with reasonable assumptions on incident rates and volunteer response distributions, as

shown in Section 4.6.3 and 4.6.4.

The MDP model presented in Section 4.3 determines which CFRs to alert and when, based on the estimated travel times of CFRs and assigned ambulances to the patient at the time of dispatch. Propositions 4.3.2 and 4.3.3 highlight key characteristics of MDP optimal policies: no additional alerts after an acceptance and immediate additional alerts if all alerted CFRs reject. Although computing the MDP optimal policies in real-time for each OHCA is unrealistic, these characteristics are useful for constructing the pre-defined policies for our selection strategy introduced in Section 4.4.

A selection strategy using a classification tree, or any classification method, can be trained using the simulation model presented in Section 4.4 to evaluate the performance of each policy for OHCA in the pre-defined set. Tables 4.1 and 4.2 show that shorter EMS response times result in fewer alerts being sent and simpler dispatch policies being optimal, given the shorter time for CFRs to respond. Figures 4.4 and 4.5 demonstrate how our tree-based adaptive selection strategy improves upon the fixed strategies currently used by CFR systems in practice.

An important observation from Section 4.6 is that the tree-selected strategy is not always desirable because it can misclassify, or it can choose to compromise patient survival for reduced volunteer fatigue. The selection strategy is most advantageous relative to the best static strategy in two extreme types of OHCA instance: when CFR response is extremely effective, leading to sending out many more alerts than the best static strategy, or when CFR response is unlikely to contribute much to patient survival, leading to refraining from sending alerts to reduce volunteer fatigue. Additionally, Section 4.6.3 highlights that volunteer acceptance rates heavily impact optimal dispatch policies and their performance.

Higher acceptance rates lead to fewer alerts being sent but improved survival and performance overall compared to lower acceptance rates.

The selection strategy approach is highly adaptive to various extensions, beyond the assumptions made herein. For example, while we assumed a fixed volunteer travel speed, with sufficient empirical data or reasonable assumptions about how volunteer travel speeds are distributed, we could evaluate the performance of each policy under random travel times. Additionally, both CFR system triage time and EMS response time significantly impact patient survival. Our simulation model can accommodate the evaluation of dispatch policies under both fixed and random triage and EMS response times. Our training for the selection strategy also allows for the inclusion of estimated EMS travel time as an additional feature to predict the appropriate policy for each incident.

There are many important directions for future research that could extend our selection strategy approach. First, we considered all volunteer arrivals after the first responder as redundant. However, high-quality CPR can be tiring, and having more than one volunteer to alternate can be beneficial. How would the optimal dispatch policies change if we start counting redundant volunteers after the first k arrivals? The tree-based selection strategy could easily be adapted to this setting by modifying how we evaluate redundant volunteers in the simulation model.

In the context of OHCA, the availability of an AED significantly impacts survival. What survival benefits could be expected if we dispatch some CFRs to grab a publicly available AED on the way to the patient? What would be considered a "good" set of policies to start with for our selection strategy in this setting? SJAS now provides small portable AEDs to a small group of active CFRs, who presumably have a higher acceptance rate compared to less engaged CFRs. Additionally,

off-duty healthcare professionals and firefighters are potentially more willing to accept an alert compared to the general public with only CPR training. How would dispatch policies change if we consider different classes of CFRs with varying acceptance probabilities and abilities to help? How do we balance the proximity of the CFR to the patient with their willingness and level of proficiency? These questions would require a model with multiple classes of CFRs, which could also be incorporated into our selection strategy approach by adding volunteer classes as predictors.

A recent podcast (Radiolab 2024) reports an overall 8% survival rate for OHCA patients in the USA, but in specific locations like casinos, the survival rate can reach as high as 54% due to a very fast response by well-trained employees. Hence, in some very special locations, there is no need to alert additional CFRs. How valuable would it be to consider the location of the OHCA as a factor in deciding the appropriate dispatch policy to use? Exploring these questions could lead to more nuanced and effective dispatch strategies, ultimately improving patient outcomes and volunteer efficiency.

APPENDIX A

CHAPTER 2 APPENDIX

A.1 Convergence to a Poisson Point Process

Consider the city as a bounded set $C \subseteq \mathbb{R}^2$. Let $(X_i : i = 1, 2, \ldots)$ be a sequence of independent random points in $C \cup \{\xi\}$, where X_i represents the location of available Community First Responder (CFR) i in the city and ξ is a dummy location representing an unavailable CFR. For each $n \geq 1$, define the point process N_n giving the (random) number of CFRs in each measurable set $B \subseteq C$ when we have a total of n CFRs as $N_n = \sum_{i=1}^{n} \epsilon_{X_i}$, where $\epsilon_{X_i}(B) = \mathbf{1}\{X_i \in B\}$. Here, B does not contain ξ, so the point process counts only *available* CFRs. Let ν_i be the location distribution of CFR i in the city, conditional on the CFR being available, and let $\alpha_i(n)$ be the probability that CFR i is available when there are a total of n registered CFRs. We take the availability probability to depend on the number of CFRs so that in the limit as $n \to \infty$ we get a nondegenerate limit. Assume that there are constants $\alpha_i \geq 0$, $i = 1, 2, \ldots$ so that for n sufficiently large, $\alpha_j(n) = \alpha_j/n$ for $j = 1, 2, \ldots, n$. Suppose there exists a finite measure μ such that for all μ-continuity sets $B \subseteq C$ (Borel sets $B \subseteq C$ whose boundary ∂B satisfies $\mu(\partial B) = 0$),

$$\lim_{n\to\infty} \sum_{i=1}^{n} \alpha_i(n)\nu_i(B) = \mu(B). \tag{A.1}$$

Proposition A.1.1. *Under the assumptions above, the point process N_n converges weakly to a Poisson point process N with mean measure μ as $n \to \infty$.*

Proof of Convergence to a Poisson Point Process

Definition A.1.2. *Let $\mathcal{B}_+$ be the set of non-negative and bounded functions on a Euclidean set E, and let P be a point process on E. For $f \in \mathcal{B}_+$ define*

$$P(f) = \int_{y \in E} f(y) P(dy) = \sum_i f(Y_i), \tag{A.2}$$

where $\{Y_i\}$ are the points of the point process P.

Definition A.1.3. *The Laplace functional of a point process P is the non-negative functional on $\mathcal{B}_+$ defined by*

$$\Psi_P(f) = \mathbb{E} \exp\{-P(f)\}. \tag{A.3}$$

Lemma A.1.4 (Resnick (2002), Theorem 4.8.2)**.** *The point process N is a Poisson point process with mean measure μ if and only if its Laplace functional is of the form*

$$\Psi_N(f) = \exp\left\{ -\int_E \left(1 - e^{-f(x)}\right) \mu(dx) \right\}, \quad f \in \mathcal{B}_+. \tag{A.4}$$

Lemma A.1.5 (Daley and Vere-Jones (2008), Proposition 11.1.VIII.)**.** *The following conditions are equivalent:*

(i) The point processes $(P_n : n \geq 0)$ converge weakly to P as $n \to \infty$.

(ii) The Laplace functional $\Psi_{P_n}(f)$ converges pointwise to the limit functional $\Psi_P(f)$ as $n \to \infty$, where f ranges over the space of continuous functions vanishing outside a bounded set.

The proof of the following lemma is straightforward and thus omitted.

Lemma A.1.6. *If $\frac{1}{n} \sum_{i=1}^n x_i \to c$, then $\lim_{n\to\infty} \sum_{i=1}^n \frac{x_i}{n - x_i} = c$.*

Proof of Proposition A.1.1. First consider $f(x) = \lambda \mathbf{1}\{x \in B\} \in \mathcal{B}_+$ where $\lambda > 0$ for some μ-continuity set $B \subseteq \mathbb{R}^2$. For n so large that $\alpha_i(n) = \alpha_i/n$ for $i = 1, 2, \ldots, n$, the Laplace functional of the point process N_n is

$$\Psi_{N_n}(f) = \mathbb{E}\left(\exp\{-N_n(f)\}\right) = \prod_{i=1}^{n} \mathbb{E}\left(\exp\left\{-f(X_i)\right\}\right) = \prod_{i=1}^{n} \left(1 - \frac{1}{n}\alpha_i \nu_i(B)(1 - e^{-\lambda})\right).$$

Define $x_i = \alpha_i \nu_i(B)(1 - e^{-\lambda})$. Any $t \in [1 - \frac{1}{n}x_i, 1]$ satisfies $1 \leq t^{-1} \leq (1 - x_i/n)^{-1}$. Integration yields

$$\frac{1}{n}x_i \leq -\ln(1 - \frac{1}{n}x_i) \leq \frac{x_i}{n - x_i}.$$

Sum over all i and exponentiate to obtain

$$\exp\left\{-\sum_{i=1}^{n} \frac{x_i}{n - x_i}\right\} \leq \Psi_{N_n}(f) \leq \exp\left\{-\frac{1}{n}\sum_{i=1}^{n} x_i\right\}$$

From (A.1), the RHS goes to $\exp\{-\mu(B)(1 - e^{-\lambda})\}$ as $n \to \infty$. By Lemma A.1.6, the LHS goes to $\exp\{-\mu(B)(1 - e^{-\lambda})\}$ as $n \to \infty$. Then by the Squeeze theorem,

$$\lim_{n\to\infty} \Psi_{N_n}(f) = \exp\{-\mu(B)(1 - e^{-\lambda})\} = \exp\left\{-\int_{\mathbb{R}^2} \left(1 - e^{-\lambda \mathbf{1}\{x \in B\}}\right) \cdot \mu(dx)\right\}$$

as required.

Now consider $f(x) = \sum_{j=1}^{k} \lambda_j \mathbf{1}\{x \in B_j\} \in \mathcal{B}_+$, in which $\lambda_j > 0, j = 1, \ldots, k$ and $B_1, \ldots, B_k$ are disjoint μ-continuity sets. The Laplace functional of the point process N_n (for sufficiently large n) is

$$\Psi_{N_n}(f) = \prod_{i=1}^{n} \mathbb{E}\left(\exp\left\{-f(X_i)\right\}\right) = 1 - \frac{1}{n}\left(\sum_{j=1}^{k} (1 - e^{-\lambda_j})\alpha_i \nu_i(B_j)\right)$$

Define $x_i = \sum_{j=1}^{k} (1 - e^{-\lambda_j})\alpha_i \nu_i(B_j)$. Then by (A.1), Lemma A.1.6 and the Squeeze Theorem, we see that

$$\lim_{n\to\infty} \Psi_{N_n}(f) = \exp\left\{-\sum_{j=1}^{k} (1 - e^{-\lambda_j})\mu(B_j)\right\} = \exp\left\{-\int_{\mathbb{R}^2} \left(1 - e^{-\sum_{j=1}^{k} \lambda_j \mathbf{1}\{x \in B_j\}}\right) \cdot \mu(dx)\right\}$$

as required.

For any general $f \in \mathcal{B}_+$, we approximate f from below with a sequence $\{f_m\} \subseteq \mathcal{B}_+$ of the form just considered, and apply the monotone and dominated convergence theorems to obtain $\lim_{n\to\infty} \Psi_{N_n}(f) = \Psi_N(f)$, where N is a Poisson point process with mean measure μ. Then, from Lemma A.1.5, N_n converges weakly to N. □

A.2 Response Time Distributions

Figure A.1 shows the cumulative distribution function (cdf) and density of the CFR response time (not counting ambulances) for several values of the expected number of available responders within range, assuming a response delay of 2 minutes and a walking pace of 0.1 km/min, and assuming that t remains small enough that the constant CFR-density assumption is valid.

A.3 CFR Density Discretization

Here we analyze the effect of a central modeling choice: that a city is split into a set of disjoint cells, in each of which CFR densities are homogeneous.

A.3.1 Data-Driven Experiment

Our discretization of a city can be interpreted as approximately capturing continuous incident locations. To see why, imagine that the location of an incident is

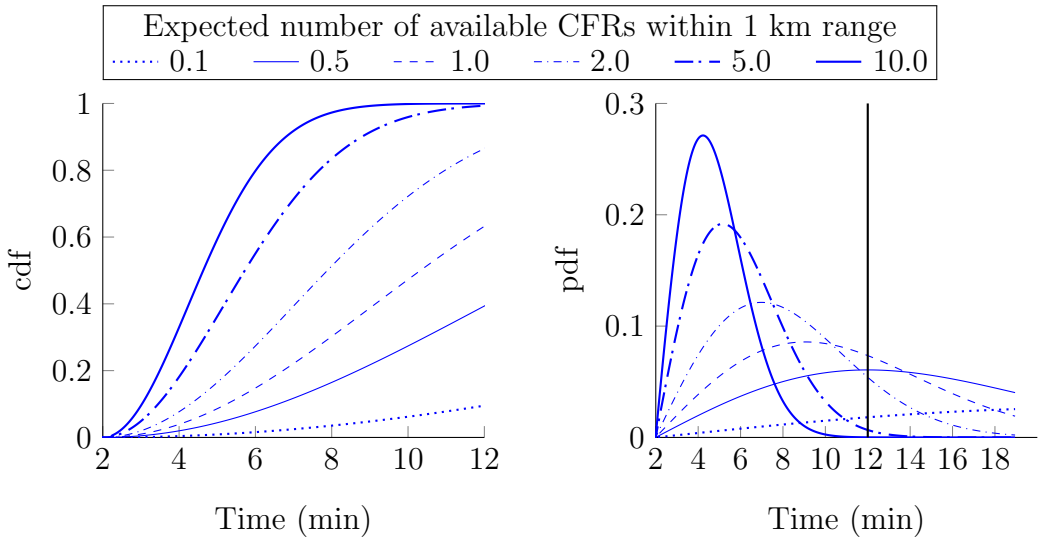


Figure A.1: Cumulative distribution function (left) and density function of on-foot response time of closest CFR, depending on the expected number of available CFRs within 1 kilometer, for a response delay κ of 2 minutes, a walking pace w of 0.1 km/min and constant CFR density. These curves are also known as Rayleigh distributions with parameter $\sigma = 1/\sqrt{2\mu\pi}$.

realized in two stages. First, we realize the (discrete) index of the region within which the incident occurred. Second, we generate the continuous location of the incident within the region. We do not explicitly perform this second step. Rather, we assume that irrespective of the location within the region, the density of CFRs around the incident location is constant. This interpretation of our model is admittedly approximate, because incident locations near the boundary of a region could receive a response from a neighboring region with a different CFR density, and it assumes that CFR density is constant throughout the region.

The impact of both of these issues on performance measures appears to be minimal. A way to appreciate why is to note an analogy with different methods to approximate a definite integral with a Riemann sum. We tested this perspective with the following experiment. We compare two scenarios:

1. As in Chapter 2, the CFR distribution has a constant density throughout a

ball of radius 1km around the incident, with a height that depends on the CFR mass allocated to the specific area unit (AU).

2. The population is modeled in more detail. Both incident locations and CFR locations arise only at the centroid of mesh blocks (MB), which are the smallest level at which Statistics NZ collects data. Auckland contains about 8600 MBs. To get a sense of scale, there are, on average, 66 MB centroids within 1km of a single MB centroid.

In both scenarios, we compute the cumulative distribution function (CDF) of the response-time distribution. These are shown in Figure A.2.

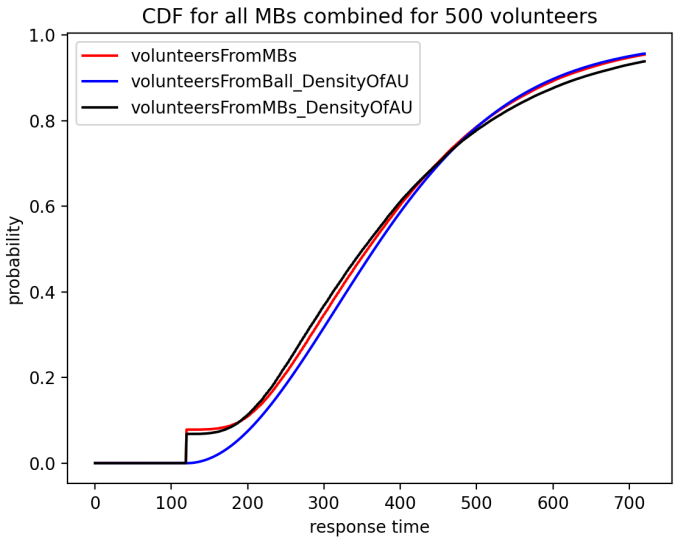


Figure A.2: Response time distribution under three different modeling assumptions.

The blue and red lines in Figure A.2 correspond to Scenarios 1 and 2 above. The difference between the two scenarios is mainly the result of a jump that can be observed in Figure A.2 in the red line at $t = 120$. This jump is caused by the location discretization of both incidents and CFRs. Because these are both assumed to be in a MB's centroid, there is an artificially high probability of having

a CFR exactly at the incident location (so we get a travel time of 0 which, combined with a response delay of 120 seconds, results in a response time of 120 seconds). That is, because CFR and incident locations follow the same discretization, part of the true probability mass is moved to $t = 120$. This effect dissipates after a response time of, say, 300 seconds. We argue that the blue line - the model from Chapter 2 - is a better representation of reality as it does not suffer from the artificial jump. Moreover, for higher response times, where we believe the red line to be a good approximation of reality (because it does not suffer from the constant density assumption and is composed of many MBs), we observe it is similar to the blue line. The black curve in the figure was used to test a third scenario: CFRs and incidents still appear in the centroid of MBs, but the volunteer masses are equal at all MBs within a single AU. This eliminates the discretization effect and singles out the constant density effect. We observe that this black line is very close to the red line described earlier, indicating that the constant density assumption only has a minor effect.

A.3.2 Thought Experiment

The experiment above suggests that, for modeling CFR response, it is not advisable to discretize a region into points, but rather to use continuous areas. In these regions, a homogeneous CFR density aids tractability, as we assume throughout Chapter 2. This raises another question: if in reality there are small differences in CFR density within one region, does a homogeneous density assumption systematically over- or underestimate the performance?

Consider a region (cell) consisting of two sub-cells. Within each sub-cell the demand and CFR densities are uniform and proportional to each other. We com-

pare two modeling approaches: consider the two sub-cells as separate, compute performance within each sub-cell and then aggregate, or 2) consider the 2 sub-cells as a single cell with uniformly distributed demand and CFR density throughout, and with the same total mass of CFRs and demand as the two sub-cells combined. In this comparison, we take late arrivals as our performance metric.

Let the population of the region be 2100 and assume 1% of them are signed up as CFRs. When splitting the region into two sub-cells, one has a population of 1000 and the other 1100. We next vary the strictness of the late arrival measure by varying the distance d_τ that a CFR can cover within the time threshold τ. For $d_\tau \geq$ 250 meters, direct calculations show that a single homogeneous cell gives better performance, while for $d_\tau < 250$ meters, splitting the region into two sub-cells gives better performance. Thus, the modeling assumption of a homogeneous CFR density throughout a region is neither always optimistic, nor always pessimistic.

A.4 Greedy Algorithm Pseudo-Code

Here we present a detailed description of the greedy algorithm, discussed in Section 2.4, that minimizes the fraction of late arrivals. The optimization problem we want to solve is

$$\begin{aligned} \min_{\nu} \quad & \sum_{l \in \mathcal{L}} \tilde{\lambda}_l e^{-\theta_l \nu_l} \\ \text{s/t} \quad & \sum_{l \in \mathcal{L}} \nu_l = 1 \\ & \nu_l \geq 0, l \in \mathcal{L}. \end{aligned} \tag{A.5}$$

The intuition behind the algorithm is that each location can be thought of as having a rate at which incidents occur that will receive a late arrival. This rate depends on the CFR mass that is already assigned to this location, as well as its

original patient rate and determines the additional allocation of CFRs.

1. Set $\nu_l = 0$ for all $l \in \mathcal{L}$.
2. Let $I = \text{argmax}\{\tilde{\lambda}_l \theta_l e^{-\theta_l \nu_l} : l \in \mathcal{L}\}$ be the index set of regions with maximal marginal benefit.
3. For all $l \in I$, increase ν_l at rate a_l until either $\sum_{l \in I} \nu_l = 1$, at which point we stop with the optimal solution, or a new region or regions joins I, i.e., $\tilde{\lambda}_l \theta_l e^{-\theta_l \nu_l} = \tilde{\lambda}_j \theta_j$ for some $j \notin I$ and any $l \in I$. (Recall that $\nu_j = 0$ at this point.)
4. Go to Step 2 above.

This greedy algorithm can be generalized to other objectives. To that end, the closed-form expression determining the point at which a new location joins the set I in Step 3 above is replaced by a step-wise evaluation of the objective. We provide the pseudo-code of the greedy algorithm for maximizing the survival objective in Algorithm 1. The evaluation function (e.g. EVALUATESURVIVAL in Algorithm 1) gives the value of the considered objective in a particular location, given a CFR distribution ν.

A.5 Example with Two Locations

Consider the problem of optimizing coverage as in Section 2.4.1. Figure A.3 depicts the optimal allocation of CFRs in a simple example consisting of two locations ($|\mathcal{L}| = 2$) of equal area. Each curve corresponds to a different value of θ_l that is common to both locations. The arrival rates, λ_1 and λ_2 at the two locations sum to 1, and we vary λ_1 on the horizontal axis. The vertical axis shows the

Data: set of all locations *Locations*, integer *nSteps*
Result: Optimal CFR distribution ν
$\nu_i \leftarrow 0$
for i **in** *Locations* **do**
  Survival[i] $\leftarrow 0$
  AddSurvival[i] $\leftarrow$ EVALUATESURVIVAL($i, \nu_i + \frac{1}{nSteps}$)
end
for $i = 1 : nSteps$ **do**
  locToAdd $\leftarrow \arg\max$(AddSurvival)
  $\nu_{locToAdd} \leftarrow \nu_{locToAdd} + \frac{1}{nSteps}$
  Survival[locToAdd] $\leftarrow$ Survival[locToAdd] + AddSurvival[locToAdd]
  AddSurvival[locToAdd] $\leftarrow$
   EVALUATESURVIVAL(locToAdd, $\nu_{locToAdd} + \frac{1}{nSteps}$) -
   Survival[locToAdd]
end
return ν
Algorithm 1: Greedy algorithm to find optimal CFR distribution ν

optimal allocation of CFRs to Location 1. The optimal allocation of CFRs to Location 1 is increasing in the fraction of calls at that location. It remains at the value 0 until the allocation to Location 2 reaches the threshold indicated in (2.4), at which point Location 1 starts to receive an allocation. The flattest of the curves corresponds to the situation with many available CFRs (large θ_l, i.e., large $n\alpha$). For this curve, the allocations to the two locations differ only modestly, because in the greedy algorithm Location 2 quickly reaches the threshold indicated in (2.4), and thereafter the two locations receive the same increase in CFR probability mass.

A.6 Frank-Wolfe Algorithm Pseudo-Code

Here we present the pseudo-code and details of the away-step Frank-Wolfe algorithm implementation for maximizing OHCA survival mentioned in Section 2.4.

The Frank-Wolfe (FW) algorithm was first introduced by Frank and Wolfe

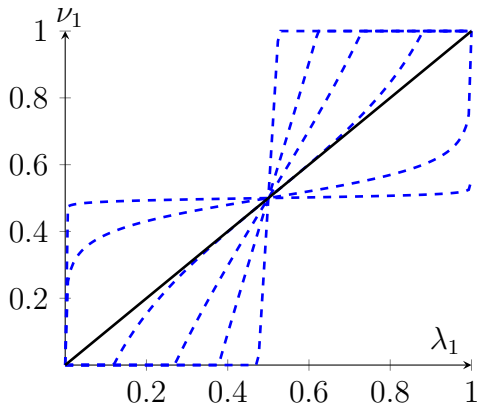


Figure A.3: Optimal distribution of CFRs over two points as a function of the fraction of demand coming from Location 1. Each curve corresponds to a single value of $\theta_l \in \{0.1, 0.5, 1, 2, 10, 100\}$ that is common to both locations. As we increase through these values, the curves approach the horizontal line at $\nu_1 = 0.5$.

(1956) and is a well-known first-order optimization method that can be applied to problems with a convex objective over a probability simplex, as is the case in our setting. The away-step FW (AFW) algorithm that we use is a modification of the original FW algorithm that avoids a problem that arises when an iterate produced by the algorithm is near the boundary of the feasible region, causing a slowing of convergence.

Lower Bounds

The major difference between our pseudo-code presented in Algorithm 2 and the AFW algorithm in Bomze et al. (2020) is the stopping criterion. In Algorithm 2, we obtain a lower bound on the optimal objective value via the cutting-plane method that relies on the convexity of the objective function. Algorithm 2 stops when the optimality gap between the objective value of the current iterate and this lower bound is smaller than some user-specified small tolerance $\delta > 0$.

Input : convex objective $f : \Delta_{d-1} \to [0,1]$, tolerance $\delta > 0$
Initialize : $x_0 \in \Delta_{d-1}, \alpha_0$
for $k = 0, 1, 2, \ldots$ **do**
  Choose $\hat{i} \in \arg\min_i \nabla_i f(x_k)$;
  $s_k \leftarrow e_{\hat{i}}$;
  $d_k^{FW} \leftarrow s_k - x_k$;
  Solve $\mathcal{P}_k$ and obtain lower bound l_k;
  if $|f(x_k) - l_k| < \delta$ **then**
    STOP;
  $S_k \leftarrow \{j : (x_k) > 0\}$;
  Choose $\hat{j} \in \arg\max_{j \in S_k} \nabla_j f(x_k)$;
  $v_k \leftarrow e_{\hat{j}}$;
  $d_k^A \leftarrow x_k - v_k$;
  if $-\nabla f(x_k)^T d_k^{FW} \geq -\nabla f(x_k)^T d_k^A$ **then**
    $d_k \leftarrow d_k^{FW}$;
    $\alpha_k^{\max} \leftarrow 1$;
  else
    $d_k \leftarrow d_k^A$;
    $\alpha_k^{\max} \leftarrow x_{\hat{j}}/(1 - x_{\hat{j}})$;
  $\alpha_k \leftarrow \min\{\alpha_k, \alpha_k^{\max}\}$;
  x_{k+1}, α_{k+1} obtained from adaptive line search subroutine;

Algorithm 2: Modified Away-step Frank-Wolfe Algorithm on Probability Simplex

Let Δ_{d-1} denote the $(k-1)$-dimensional probability simplex,

$$\Delta_{d-1} = \{x \in \mathbb{R}^d : \sum_{i=1}^{d} x_i = 1, x \geq 0\}.$$

Let $f : \Delta_{d-1} \to [0,1]$ be the death probability function, which is convex according to Propositions 2.4.2 and 2.4.3. We assume that f is differentiable, which holds for the coverage and survival functions we used in our case study. By convexity, for all $x, y \in \Delta_{d-1}$,

$$f(y) \geq f(x) + \nabla f(x)^T (y - x).$$

Given the sequence of iterates $\{x_k\}_{k=1}^n \subseteq \Delta_{d-1}$ produced by Algorithm 2 when minimizing $f(x)$ over the probability simplex, we can obtain a lower bound on the

objective value by solving the linear program with decision variables y and z,

$$\begin{aligned} \min \quad & z \\ \text{s.t.} \quad & z \geq f(x_k) + \nabla f(x_k)^T (y - x_k) \quad \forall k = 1, \ldots, n \\ & y \in \Delta_{d-1} \end{aligned} \tag{$\mathcal{P}_n$}$$

In each iteration of Algorithm 2, we compute the gradient $\nabla f(x_k)$ of the new iterate x_k, add an additional constraint (cut) to $\mathcal{P}_n$ and update the lower bound by solving $\mathcal{P}_n$.

Adaptive Line Search Algorithm

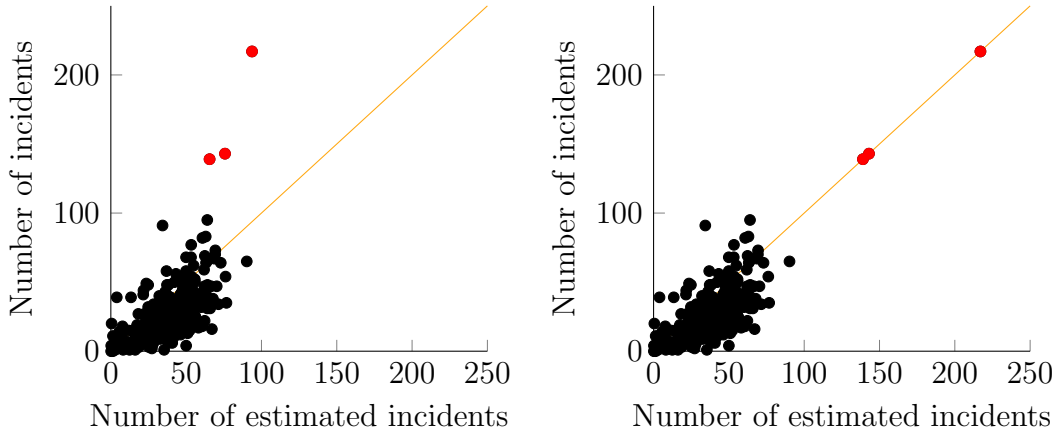
We use a line search algorithm that is similar to the one proposed by Berahas et al. (2021). Let $\theta, \gamma \in (0, 1)$ be fixed constants. In each iteration k, let x_k be the current iterate, and let d_k denote the current search direction. We start the line search with a step size parameter α_k obtained using the previous step size as detailed below. Consider $x_k^+ = x_k + \alpha_k d_k$. The current step size is chosen to satisfy the sufficient decrease Armijo condition

$$f(x_k^+) \leq f(x_k) + \alpha_k \theta \nabla f(x_k)^T d. \tag{A.6}$$

If the Armijo condition (A.6) is satisfied, we proceed to the next iteration with $x_{k+1} = x_k^+$ and $\alpha_{k+1} = \gamma^{-1} \alpha_k$. Otherwise, we keep updating the current step size $\alpha_k \leftarrow \gamma \alpha_k$ until (A.6) is satisfied.

A.7 Estimating the OHCA Incidence Rate

In this section, we provide more details on how we estimate the OHCA incidence rate. We use four socio-economic factors (Maori, Pacific, male and elderly population) that are identified as significantly affecting the OHCA rates according to Table 1 in Dicker et al. (2019). Each 1% increase in the Maori and Pacific population relative to a baseline average across all area units gives a 1% increase in the OHCA rate, and each 1% increase in the male and elderly population leads to a 3.7% increase in the OHCA rate. We compute a baseline OHCA rate using the data provided by the SJAS. We then adjust the incidence rate in each area unit depending on how the area unit's socio-economic factors deviate from their baseline average. The final estimated count is then obtained by multiplying by the population in each area unit, because Dicker et al. (2019) estimates the OHCA rate per person year. The socio-economic data as well as the population in each area unit are given in the 2013 census (Stats New Zealand Geographic Data Service 2015b).



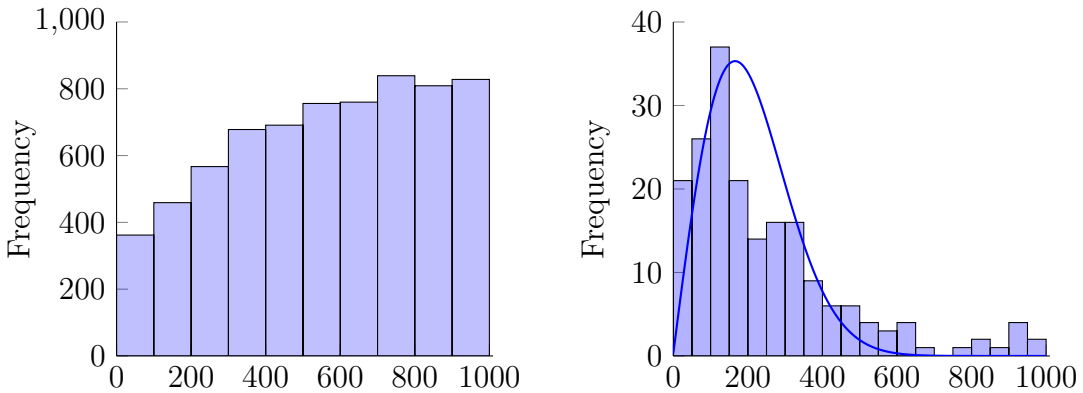
(a) Empirical vs Estimated Incidents (b) Empirical vs Estimated Incidents

Figure A.4: Scatter plot of the actual number of cardiac arrests and the estimated number of cardiac arrests between 2013 and 2020 over all area units before (left) and after (right) adjusting the CBD and the Airport.

Figure A.4 shows the number of OHCA incidents since 2013 versus the number of estimated OHCA incidents using Dicker et al. (2019) before and after a manual adjustment of the estimates for the CBD (two area units) and the airport (one area unit). We see quite a bit of dispersion around a general linear trend. Some of the overestimation is perhaps due to the fact that Dicker et al. (2019) estimates the OHCA rate per person year, ignoring any influx of people. The three area units highlighted in Figure A.4, which are the CBD and the area unit containing the airport, have a very large influx from people working and traveling there. As a result, we can see from Figure A.4a that the actual number of OHCAs in these 3 outlier area units is greatly underestimated. We therefore replace the estimated OHCA counts of these 3 area units with empirical counts from SJAS data. The OHCA counts in these 3 area units are reasonably high which suggests that using the empirical counts should yield modest estimation error relative to the true values in those area units.

A.8 Retrospective Analysis

Here we compare predictions from the model in Chapter 2 to observed data in the GoodSAM system in Auckland, New Zealand. Figure A.5a provides a histogram of the distances between incidents and alerted CFRs over the 6749 alerts in the data. The corresponding shape for a Poisson point process with constant intensity is triangular if *all* CFRs within 1 kilometer are alerted. The data has more mass near shorter distances than a triangle, which is expected since the subset of CFRs that were alerted are those closest to an incident. The location of *all* CFRs at the time of an incident is not recorded due to privacy concerns.



(a) Alerted CFRs

(b) Closest CFRs

Figure A.5: Histograms of distances from incidents to CFRs. Left: Distances of alerted CFRs from the 6749 alerts in Auckland since December 2017. Right: Distance of the *closest* alerted CFR for 194 incidents in the Auckland CBD since December 2017. The overlaid curve is the theoretical density function that follows from our model.

Now consider the distribution of distance to the *closest* CFR. We cannot easily combine predictions from area units l with different CFR mass ν_l because the peak of the distribution varies; we instead focus on two area units in the central business district (CBD) with relatively high incident rates: Figure A.5b provides the corresponding histogram. Also depicted is the density of the distance to the closest CFR for a Poisson process with a mean of 12 available CFRs in that area. The value of 12 was chosen because it yielded a reasonable fit. The fit is at least plausible, though certainly far from perfect. It should be noted that the data consists of all alerts from inception of the program, and so it includes periods with varying numbers of CFRs, degrading the potential fit.

A.9 Profile Sensitivity Analysis

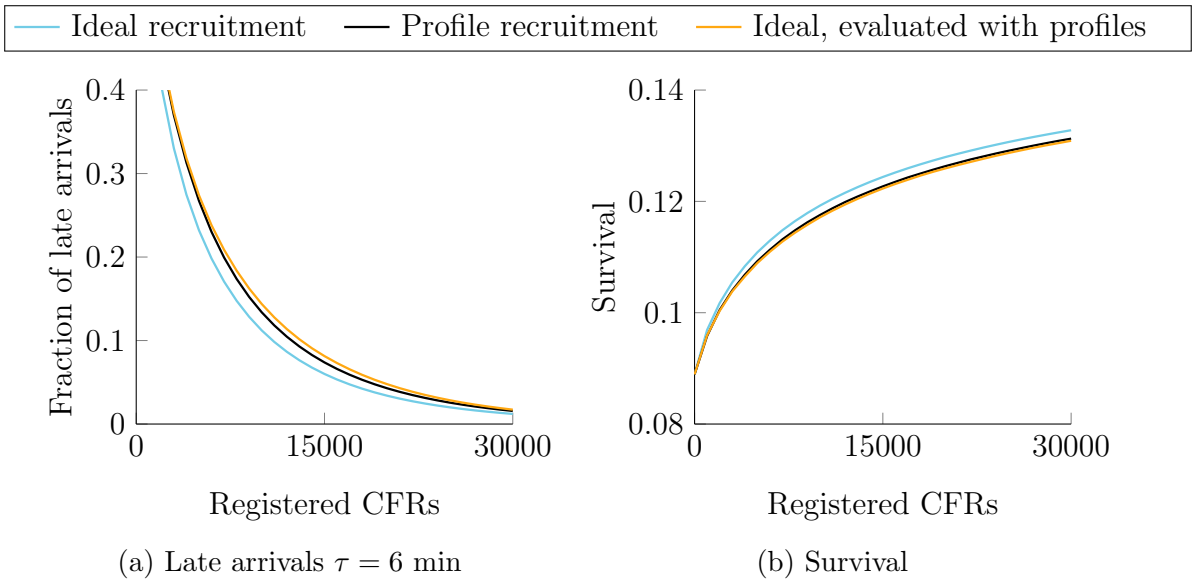
As discussed in Section 2.5.3, our profile estimates are not perfect because we are unable to access reliable and fine-grained mobility data in Auckland. In this section, we consider the impact of profile estimation on the results through a sensitivity analysis.

A.9.1 Impact of Profile

First, we explore the impact of profile recruitment relative to recruitment when ignoring profiles. Figure A.6 shows the impact of profiles on two different objectives. For the late-arrivals objective, profile recruitment has an appreciable impact; there is a nontrivial performance gap with ideal recruitment. Moreover, if one recruits CFRs using the ideal-recruitment allocation, then the realized performance will be much worse than predicted if the CFR profiles represent reality. For the survival objective, the difference in performance between ideal and profile recruitment is small. Still, even in these cases, there is an unnecessary reduction in performance if one ignores profiles when recruiting.

A.9.2 Uncertainty Analysis

Second, we perform uncertainty analysis in which we assume that we optimize under the estimated profiles, while true profiles are "noisy" versions of the estimated ones. More specifically, considering the estimated profile p_l for area unit l, we construct $\hat{p}_l$ such that for all area units $i \in \mathcal{L}$, $\hat{p}_{li} = p_{li}$ if $p_{li} > 0$, otherwise $\hat{p}_{li} = \epsilon$ for some arbitrarily small $\epsilon > 0$. We then assume the true profiles of each area



(a) Late arrivals $\tau = 6$ min

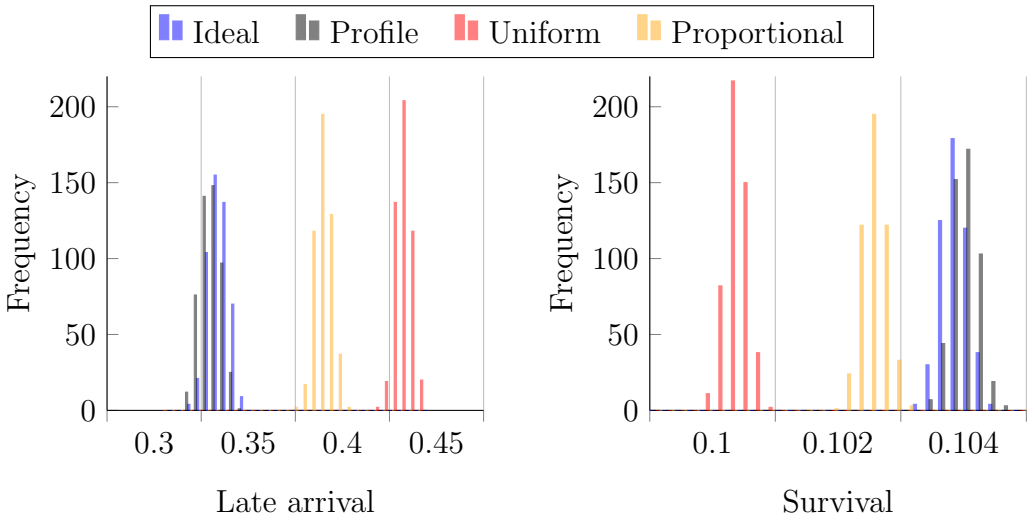
(b) Survival

Figure A.6: Impact of an increasing number of available CFRs for different CFR distributions (ν) on the fraction of late arrivals ($\tau = 6$ min) and the survival probability with EMS response for the optimal CFR distributions (ν) with and without profiles. Note that (a) and (b) have different scales.

unit l follow a Dirichlet distribution with parameters $\hat{p}_l$. We again illustrate our results with both late arrival probability and the Waalewijn survival function and 500 expected available CFRs.

Figure A.7 suggests that our solutions are reasonably robust to the noise in profile estimation as long as the estimated profiles are not too far from the true ones. In addition, it shows that recruitment based on our optimization model, whether with or without profiles, indeed yields better performance in both objectives compared to other reasonable recruitment strategies. For the late-arrival objective, both optimal recruitment strategies bring down the late arrival from 45% to approximately 35%. The survival probability increases from 10.1% to approximately 10.5%, which in Auckland is equivalent to 12 more lives saved annually.

In these results, the performance of optimized solutions with or without profiles is almost indistinguishable, which suggests that when an accurate profile estimation



(a) Late arrivals $\tau = 6$ min

(b) Survival

Figure A.7: Impact of inaccurate profile estimation for different recruitment distributions on the fraction of late arrivals ($\tau = 6$ min) and the survival probability with EMS response for the optimal CFR distributions (ν) with and without profiles.

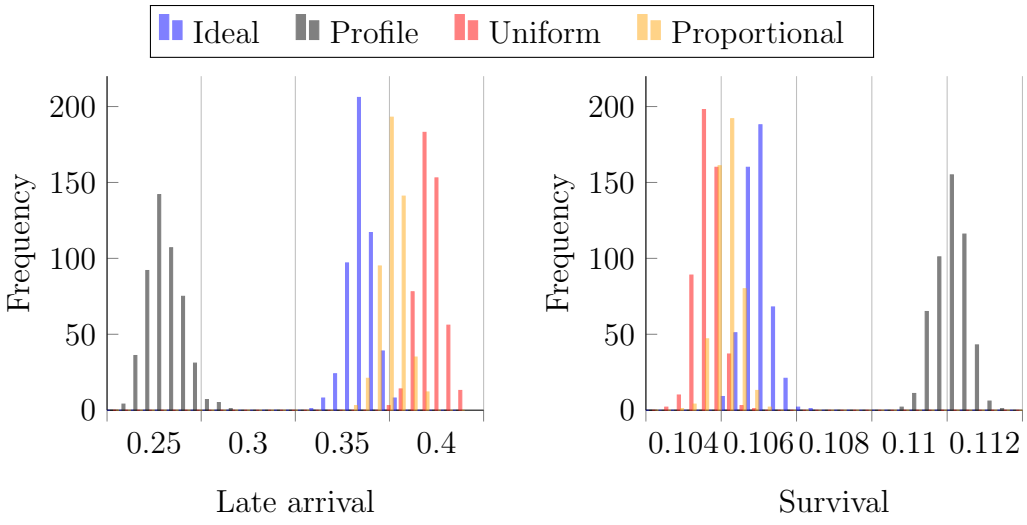
is not possible, optimization under the assumption that CFRs always stay in their home location could give good performance. The small difference is possibly due to the fact that the majority of people are still modeled to stay at home in our estimated profile either because they do not work, or because they work near their home location.

A.9.3 Example with Larger Profile Impact

Here we provide an additional example with a different demand distribution and different profiles than we used in our case study to show that considering profiles could lead to a larger impact on performance metrics.

We first generate a random demand distribution using the Dirichlet distribution, Dir(λ), where λ is the input demand distribution to our case study. The

random profile is then obtained by randomly shuffling the rows of our estimated profile matrix. We then perform uncertainty analysis just as above with the new demand distribution and profile.



(a) Late arrivals $\tau = 6$ min

(b) Survival

Figure A.8: Impact of inaccurate profile estimation for different recruitment distributions on the fraction of late arrivals ($\tau = 6$ min) and the survival probability with Fixed EMS response time for the optimal CFR distributions (ν) with and without profiles.

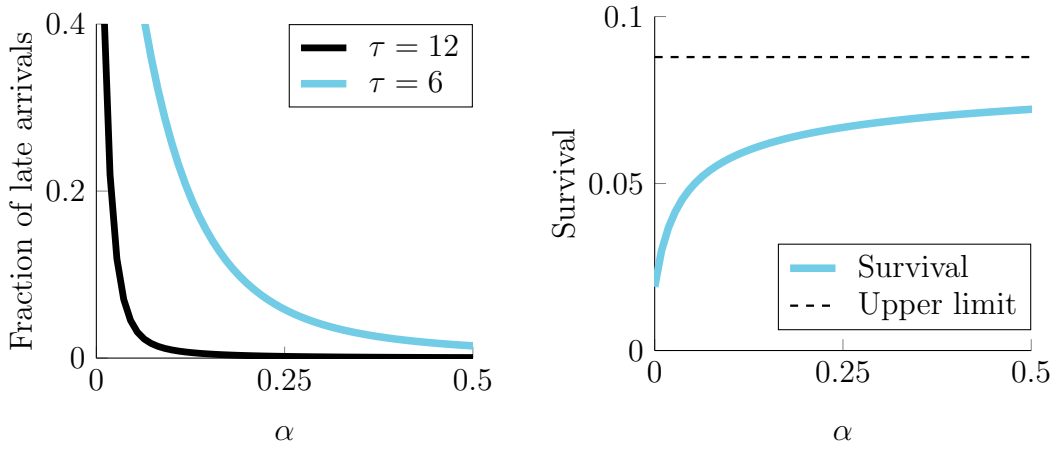
Figure A.8 shows that recruitment based on optimization either with or without profiles still leads to better performance in both objectives compared to other recruitment strategies, while the robustness of the solutions remains. However, the performance of optimal recruitment with profiles is much better than optimal recruitment without profiles; the late arrival decreases from 38% to about 27%, and the survival probability increases from about 10.5% to 11.1%. This suggests that considering profiles could lead to a significant difference in CFR performance in some settings.

A.10 Homogeneous Ambulance Response

The effect of CFRs is somewhat masked by the presence of ambulances. To see the benefits of CFRs more directly, we evaluate late arrival without ambulances and evaluate survival by considering a scenario with a homogeneous ambulance response of 12 minutes throughout the Auckland region.

A.10.1 Quantifying the Benefit of a Given CFR Base

We quantify the benefit of a given CFR base as detailed in Section 2.6.2 with the acceptance rate $\alpha \in [0, 0.5]$.



(a) Late arrivals (b) Survival

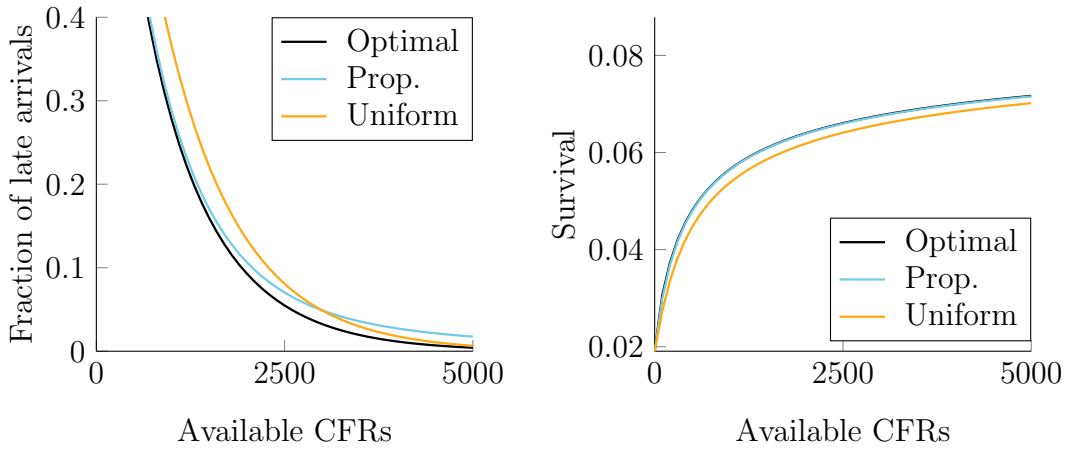
Figure A.9: Impact of increasing the CFR availability α on fraction of late arrivals and survival, assuming 1% of the population is signed up as CFR, distributed proportionally over the area units.

Figure A.9a shows that in the case of no ambulances, bringing down response times that exceed the 12-minute target to 1% requires an α of 10% (1100 available CFRs). While not depicted in the figure, to satisfy the requirement of reaching 50% of the incidents within 6 minutes, $\alpha \approx 4.6\%$ suffices. In Figure A.9b, we see

that under fixed 12-minute ambulance responses, the upper limit of survival is only about 8.8%. Nonetheless, CFRs still contribute substantially to survival even with low values of α. Increasing the survival probability from 1.9% (with no CFRs) to 6% requires an alpha of 13.6% (about 1500 available CFRs).

A.10.2 Quantifying the Impact of CFRs' Geographical Spread

As in Section 2.6.3, we consider two scenarios: one where CFRs are uniformly distributed over the city and another where CFRs are distributed proportionally to demand, and then bound the performance using the optimal CFR distributions. Compared to Figure 2.3a, it is clearer from Figure A.10a that with a small number of available CFRs, the performance of a proportional allocation is indistinguishable from that of the optimal CFR distribution.

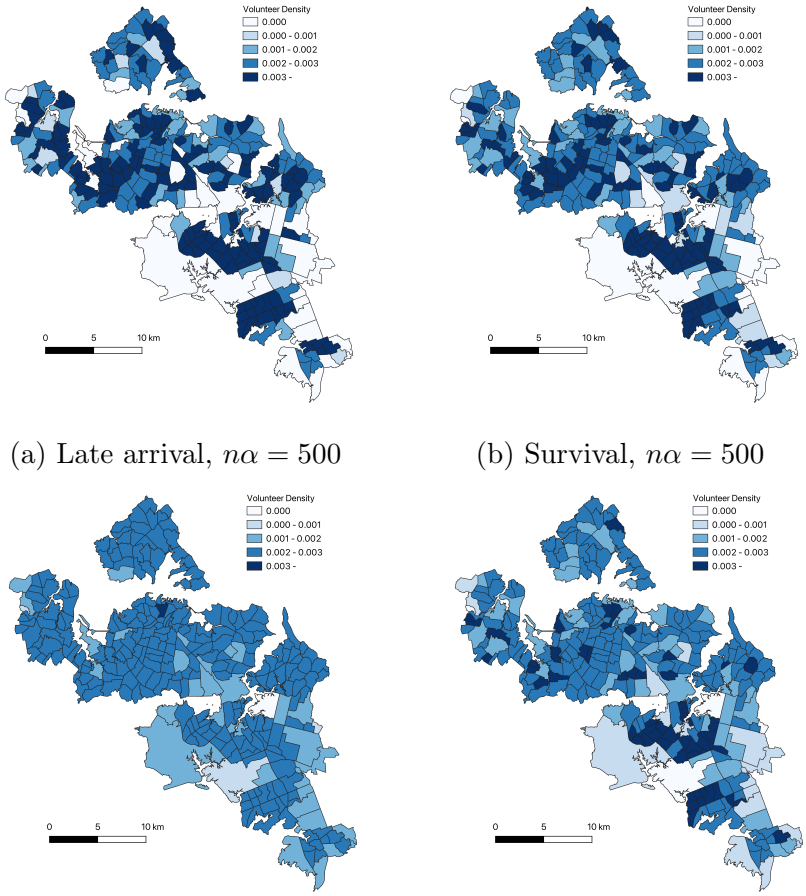


(a) Late arrivals $\tau = 6$ min

(b) Survival

Figure A.10: Impact of increasing number of available CFRs ($n\alpha$) for different CFR distributions (ν) on the fraction of late arrivals ($\tau = 6$ min) and the survival probability. The black lines are each formed by the solution of 50 optimization problems. Note that (a) and (b) have different scales.

Figure A.11 compares the geographical distribution of an optimal ν for both a limited number of CFRs ($n\alpha = 500$) and a larger number of CFRs ($n\alpha = 5000$) under two objectives (late arrivals with no ambulances and patient survival with fixed 12 minute ambulance response time). Without the influence of ambulances, Figure A.11c shows more clearly that a distribution of CFRs with a higher number of CFRs resembles a uniform distribution.



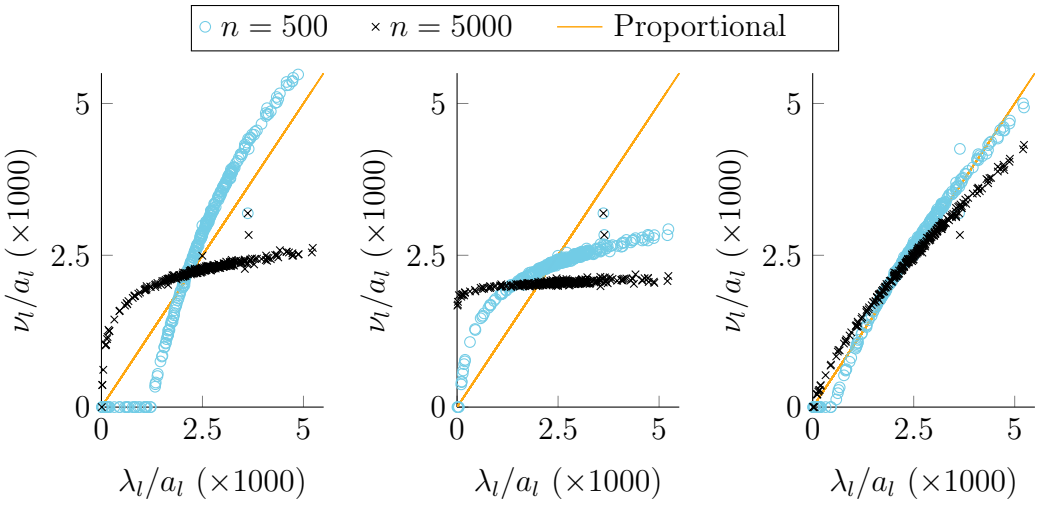
(a) Late arrival, $n\alpha = 500$ (b) Survival, $n\alpha = 500$

(c) Late arrival, $n\alpha = 5000$ (d) Survival, $n\alpha = 5000$

Figure A.11: Optimal CFR distribution (fraction of CFRs in that region divided by surface area) to minimize late arrivals for $\tau = 6$ min (a and c) or to maximize survival (b and d), with $n\alpha = 500$ or $n\alpha = 5000$. Darker colors represent higher CFR density.

A.11 Ideal Recruitment

Figure A.12 plots the optimal CFR density in each area unit as a function of the population density. For comparison, a solution where CFR densities are proportional to demand is also shown.



(a) Late arrival $\tau = 6$ (b) Late arrival $\tau = 12$ (c) Survival

Figure A.12: Optimal allocations of probability mass for late arrivals with different threshold times and survival, contrasted with a distribution of CFRs proportional to the demand density. Each mark represents one region (area unit): the horizontal axis shows the demand density per km^2 in that region; the vertical axis shows the fraction of CFRs in that region divided by its area.

With only 500 available CFRs, the optimal solution for late arrivals has no CFRs in areas with low demand density. Moreover, a comparison of Figure A.12a and A.12b shows that the demand density above which a region receives CFRs is smaller when the response-time target is bigger. The curves for 5000 available CFRs show that in this scenario hardly any region is left without CFRs regardless of the threshold. This behavior is in line with the insights obtained from (2.4). Furthermore, in Figure A.12b we see that for a large number of CFRs with a high response-time target, the optimal ν is almost uniformly distributed.

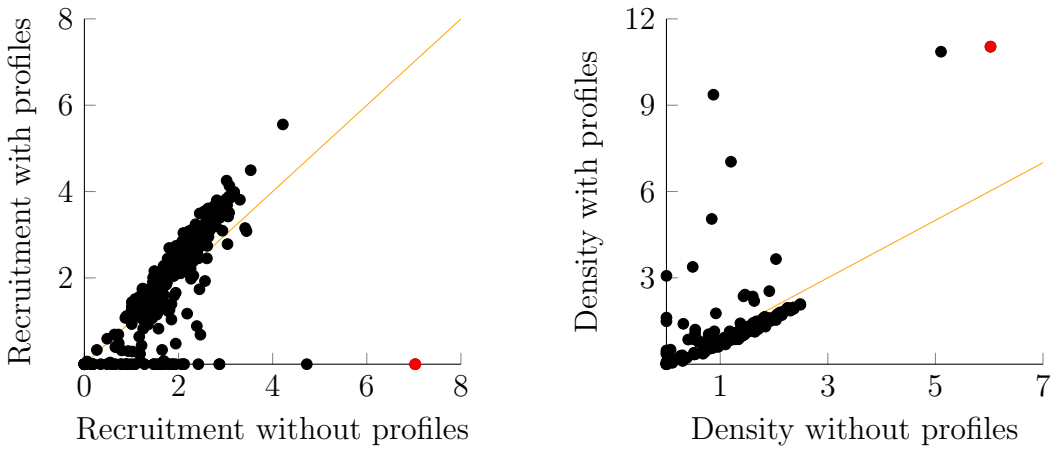
The plot containing the optimal ν for survival (Figure A.12c) shows that optimal solutions are rather similar for a wide range of CFRs (500-5000). Moreover, both solutions appear close to a ν that is proportional to the demand density. We do see, however, that for a small number of CFRs the focus is slightly more on high-demand areas, leaving some regions with zero CFRs. For a high number of CFRs, low-demand areas receive slightly more CFRs than a proportional allocation would assign to them.

The plots in Figure A.12 exhibit a few outlier points. These arise due to numerical issues in numerical optimization, especially with regions of very small areas, e.g., one region has an area of less than 0.25km^2.

A.12 Profile Recruitment

To isolate the effect of profiles, we consider a scenario with a single time period and a homogeneous ambulance response throughout the region. We look at the optimal recruitment per area unit for the case of 500 available CFRs and the survival objective. Figure A.13 shows recruitment effort and the resulting distribution ν under these assumptions, where each point represents a single area unit.

In Figure A.13a, both axes represent the number of available CFRs recruited for each area unit. For example, the rightmost point (highlighted in red) shows that we desire about 7 available CFRs (that is, $7/0.14 \approx 50$ registered CFRs) in the area unit belonging to that point. This happens to be the CBD. The vertical axis shows that for this same area unit, under profiles, no recruitment is necessary. The reason is that this area has a sufficient influx of CFRs from other area units, as is the case for all points on the horizontal axis. Most points, however, are slightly



(a) Recruitment difference (b) Under and over-served areas

Figure A.13: Scatter plots of (a) the optimal number of available CFRs recruited and (b) the resulting CFR densities (available CFRs per square km), with and without profiles, when maximizing survival for 500 available CFRs. Each point represents one area unit and in both figures, the CBD is highlighted in red.

above the diagonal in an apparently linear way. These are points that need more recruitment to compensate for an outflux of CFRs.

Figure A.13b compares the resulting CFR density ν under optimal recruitment in situations with and without profiles. For example, the red point, which again represents the CBD, indicates that we want about 6 available CFRs per km^2 there (under the restriction that the total is 500 city-wide). However, despite not recruiting any CFRs who live there, the large influx of workers results in about 11 available CFRs per km^2. Similarly, all points above the diagonal are over-resourced due to their influx. As a result of this over-serving, most other area units lie below the diagonal and thus have a slightly lower CFR density compared to the ideal. As a result of this over-resourcing of certain areas at the cost of others, the overall survival rate decreases from 4.82% to 4.62%.

APPENDIX B

CHAPTER 4 APPENDIX

B.1 Proof for MDP Optimal Policies

In this appendix, we provide the proofs for the propositions presented in Section 4.3, which describe important characteristics of the MDP optimal policies. We then present another interesting property of the MDP formulation when volunteer fatigue is defined differently.

Proposition 4.3.1. *The Pareto-optimal frontier is concave with respect to the expected volunteer fatigue for any fixed OHCA.*

Proof. Consider any single OHCA. Let π_1, π_2 be two policies with $(g_{\pi_1}, f_{\pi_1}), (g_{\pi_2}, f_{\pi_2})$ being the expected volunteer fatigue and expected survival probability for the two policies π_1, π_2 respectively. For any $\lambda \in [0, 1]$, consider the randomized policy π_λ that dispatches under policy π_1 with probability λ and dispatches under policy π_2 with probability $1 - \lambda$, independently of all else. Under policy π_λ, the expected volunteer fatigue is $\lambda g_{\pi_1} + (1 - \lambda) g_{\pi_2}$, and the expected survival probability is $\lambda f_{\pi_1} + (1 - \lambda) f_{\pi_2}$. Thus, the set $G = \{(g, f) : (g, f) \text{ is attained by a policy}\}$ is convex. The result follows immediately. □

Proposition 4.3.2. *Once someone accepts, the optimal policy π^* will not send out any additional alerts.*

Proof. Consider a state $s_t = (t, \mathbf{v}_t, r_t)$, where m volunteers have been alerted, and at least one volunteer has accepted the alert. Let π be a policy that issues an additional alert in state s_t. Contrast this with a policy π' that sends no further

alerts in state s_t from time t onwards. The values of s_t under policies π and π', using (4.2), are

$$V^{\pi}(s_t) = f_{\pi}(s_t) - \gamma g_{\pi}(s_t), \text{ and}$$
$$V^{\pi'}(s_t) = f_{\pi'}(s_t) - \gamma g_{\pi'}(s_t),$$

where $f_{\pi}(s_t)$ and $f_{\pi'}(s_t)$ represent the expected survival probability, and $g_{\pi}(s_t)$ and $g_{\pi'}(s_t)$ represent the expected volunteer fatigue given the state s_t following policies π, π' respectively.

The expected survival under policies π and π' remains the same because the additional volunteer alerted by π is further away than the responding volunteer. However, π sends more alerts, resulting in a higher number of expected redundant volunteers ($g_{\pi}(s_t) \geq g_{\pi'}(s_t)$). As a result, $V^{\pi}(s_t) \leq V^{\pi'}(s_t)$, indicating that π is never optimal. □

Proposition 4.3.3. *If all alerted volunteers have rejected alerts by time t and, under the optimal policy π^*, no new alert is sent at time t, then no alerts will be sent out afterward.*

Proof. Let $\mathbf{v}$ be a volunteer status vector such that all m alerted volunteers have rejected the alert. Consider $s_t = (t, \mathbf{v}, r)$, a non-terminal state where the optimal policy does not send out any additional alert in state s_t, i.e. $\pi^*(s_t) = 0$.

We consider the potential next states for different actions. If $a_t = 0$ (no alert is sent), then s_t transitions to state $s_{t+1} = (t+1, \mathbf{v}, r-1)$ with probability 1. If, instead, $a_t = 1$ (the next closest volunteer is alerted) then, letting $\mathbf{v}'$ be the volunteer status vector updated to reflect the newly alerted volunteer, the system transitions to state $s'_{t+1} = (t+1, \mathbf{v}', r-1)$ with probability 1.

Given that the optimal policy does not send out any alerts in state s_t, we know

$$V^*(s_t) = V^*(s_{t+1}) > V^*(s'_{t+1}). \tag{B.1}$$

Now suppose the optimal policy π^* issues an additional alert in the state s_{t+1}, i.e. $\pi^*(s_{t+1}) = 1$. Then, the only possible next state of s_{t+1} is $s'_{t+2} = (t+2, \mathbf{v}', r-2)$, and

$$V^*(s_{t+1}) = V^*(s'_{t+2}) \leq V^*(s'_{t+1}).$$

The inequality follows directly from Proposition B.1.1. However, this contradicts with Equation (B.1) above. Thus it is not optimal to issue a new alert at time $t+1$. A straightforward inductive argument completes the proof. □

So far, we have defined volunteer fatigue as proportional to the number of redundant volunteers on-scene. However, multiple factors could contribute to volunteer fatigue. Another possible factor is when volunteers receive too many alerts, leading to annoyance and disengagement. Therefore, penalizing the number of alerts sent can also be a way of defining volunteer fatigue. Under this definition, we obtain the following interesting property of the MDP formulation.

Proposition B.1.1 indicates that there are diminishing returns when the status of volunteers remains unchanged over time, as this implies volunteers arrive later, which in turn results in a lower survival probability.

Proposition B.1.1. *If volunteer fatigue is defined as proportional to the number of alerts sent, then when volunteer status remains the same, as time goes on, the value of states decreases, i.e. for any non-terminal states* $s_t = (t, \mathbf{v_t}, r_t)$ *and* $s_{t+1} = (t+1, \mathbf{v_{t+1}}, r_{t+1} = r_t - 1)$*, we have* $V^*(s_t) \geq V^*(s_{t+1})$.

Before we prove Proposition B.1.1, we first present the following lemma, which shows that the diminishing return holds even when someone just accepts the alert and the remaining fastest responder travel time becomes the same.

Lemma B.1.2. *Consider volunteer status vector* $\mathbf{v}$ *such that at least one of the alerted volunteers has accepted the alert. For any non-terminal states* $s_1 = (t, \mathbf{v}, r)$ *and* $s_2 = (t+1, \mathbf{v}, r)$*, we have* $V^*(s_1) \geq V^*(s_2)$*.*

Proof. We consider induction on the remaining travel time steps of the fastest responder r. When $r = 1$, all possible next states of s_1 and s_2 are terminal. By Proposition 4.3.2, no alerts will be sent at states s_1 and s_2 ($\pi^*(s_1) = \pi^*(s_2) = 0$). Let m be the total number of volunteers alerted in $\mathbf{v}$. Then the values of s_1 and s_2 under the optimal policy π^*, using (4.4), are

$$V^*(s_1) = \phi(t_{CPR,1}, t_{EMS}) - \gamma m, \text{ and}$$
$$V^*(s_2) = \phi(t_{CPR,2}, t_{EMS}) - \gamma m,$$

where $\phi(t_{CPR,1}, t_{EMS})$ and $\phi(t_{CPR,2}, t_{EMS})$ gives the expected survival rate of s_1, s_2 using (4.1), where t_{EMS} is fixed for both states, and $t_{CPR,1}$ and $t_{CPR,2}$ are the minimum of the sums of $t+1$ and $t+2$ with the witness delay and t_{EMS}, respectively. π^* gives a higher expected survival rate for s_1 following (4.1), because s_1 is one time step prior to s_2 ($t_{CPR,1} \leq t_{CPR,2}$) and thus has a higher probability of reaching the patient faster with the same first responder travel time. As a result, $V^*(s_1) \geq V^*(s_2)$.

Now assume Lemma B.1.2 holds for $1, \ldots, r-1$. Similarly, we know from Proposition 4.3.2 that $\pi^*(s_1) = \pi^*(s_2) = 0$. The values of s_1 and s_2 under the

optimal policy π^*, using (4.4), are

$$V^*(s_1) = \sum_{s_1'} p(s_1'|s_1, 0) \left[h(s_1, 0, s_1') + V^*(s_1')\right], \text{ and}$$
$$V^*(s_2) = \sum_{s_2'} p(s_2'|s_2, 0) \left[h(s_2, 0, s_2') + V^*(s_2')\right].$$

Consider all possible evolutions $\mathbf{v}'$ from the volunteer status vector $\mathbf{v}$, given that there is no alert sent out at states s_1 and s_2, there is a one-to-one correspondence between s_1' and s_2' under the same volunteer status vector $\mathbf{v}'$. There are two scenarios to consider:

1. *No new volunteer acceptance that improves upon the fastest responder* : This scenario gives us $s_1' = (t+1, \mathbf{v}', r-1)$ and $s_2' = (t+2, \mathbf{v}', r-1)$.
2. *New acceptance improves upon the fastest responder* : Without loss of generality, assume it is volunteer n with $d_n < r-1$, which gives us $s_1' = (t+1, \mathbf{v}', d_n)$ and $s_2' = (t+2, \mathbf{v}', d_n)$.

In both scenarios, s_1' and s_2' are non-terminal, so no reward are accrued ($h(s_1, 0, s_1') = h(s_2, 0, s_2') = 0$) according to (4.2). Since $V^*(s_1') \geq V^*(s_2')$ follows immediately from the induction hypothesis, we have $V^*(s_1) \geq V^*(s_2)$. A straightforward inductive argument completes the proof. □

Proof of Proposition B.1.1. Similarly, we consider induction on the remaining travel time steps of the fastest responder r_t in state s_t. We start with $r_t = 2$ to ensure s_{t+1} is non-terminal with $r_{t+1} = 1$. Note that the optimal policy π^* will not send any alerts when $r = 1$, because even if the alerted volunteer responds immediately, they will not improve upon the fastest responder who has arrived by the time of the response.

All possible next states of s_{t+1} are terminal states. Let m be the total number of volunteers alerted in $\mathbf{v}$. The value of $s_{t+1} = (t+1, \mathbf{v}, 1)$, using (4.4), is given by

$$V^*(s_{t+1}) = \sum_{s'} p(s'|s_{t+1}, 0) \left[h(s_{t+1}, 0, s') + V^*(s')\right] = \phi(t_{CPR}, t_{EMS}) - \gamma m,$$

where t_{CPR} is the minimum of the sum of $t+2$ with the witness delay and t_{EMS}.

Now consider the value of state s_t if no alert is sent at s_t ($a_t = 0$),

$$Q(s_t, a_t = 0) = \sum_{s'} p(s'|s_t, 0) \left[h(s_t, 0, s') + V^*(s')\right].$$

We know $Q(s_t, a_t = 0) \leq V^*(s_t)$ following the definition of the optimal policy in (4.5). If $\mathbf{v}$ has no outstanding alerts (alerts with no response), then the only possible next state of s_t is s_{t+1} and $Q(s_t, a_t = 0) = V^*(s_{t+1})$. Even if there exist outstanding alerts, they will not improve upon the response time as we discussed above, i.e., $s' = (t+1, \mathbf{v}', 1)$. Then $Q(s_t, a_t = 0) = V^*(s_{t+1})$ because all possible evolution $\mathbf{v}'$ of the volunteer status vector $\mathbf{v}$ has m alerts sent out under the optimal policy π^* and all possible next states s' and s_{t+1} gives the same volunteer response time. As a result, $V^*(s_t) \geq Q(s_t, a_t = 0) = V^*(s_{t+1})$.

Now assume Proposition B.1.1 holds for $2, \ldots, r_t - 1$. Denote the action taken by the optimal policy π^* at state s_{t+1} as $a^*_{t+1} = \pi^*(s_{t+1})$. The value of s_t following action a^*_{t+1} is no larger than the value of s_t under the optimal policy π^*, i.e., $Q(s_t, a_t = a^*_{t+1}) \leq V^*(s_t)$ following (4.4) and (4.5):

$$Q(s_t, a_t = a^*_{t+1}) = \sum_{s'_t} p(s'_t|s_t, a^*_{t+1}) \left[h(s_t, a^*_{t+1}, s'_t) + V^*(s'_t)\right], \text{ and}$$

$$V^*(s_{t+1}) = \sum_{s'_{t+1}} p(s'_{t+1}|s_{t+1}, a^*_{t+1}) \left[h(s_{t+1}, a^*_{t+1}, s'_{t+1}) + V^*(s'_{t+1})\right].$$

Consider all volunteer evolutions $\mathbf{v}'$ from $\mathbf{v}$ following the action the a^*_{t+1}. This only depends on the m alerted volunteers in $\mathbf{v}$ as a^*_{t+1} only decides whether the

$m+1$-th closest volunteer is alerted or not. Similarly, there is a one-to-one correspondence between s'_t and s'_{t+1}, the possible next states of s_t and s_{t+1}, respectively, if they have the same volunteer status vector $\mathbf{v}'$. There are two scenarios to consider:

1. *No new volunteer acceptance that improves upon the fastest responder* : This scenario gives us $s'_t = (t+1, \mathbf{v}', r_t - 1)$ and $s'_{t+1} = (t+2, \mathbf{v}', r_t - 2)$.
2. *New acceptance improves upon the fastest responder* : Without loss of generality, assume it is volunteer n with $d_n \leq r_t - 2$, which gives us $s'_t = (t+1, \mathbf{v}', d_n)$ and $s'_{t+1} = (t+2, \mathbf{v}', d_n)$.

Under both scenarios, s'_t and s'_{t+1} are non-terminal, so no reward are accrued ($h(s_t, a^*_{t+1}, s'_t) = h(s_{t+1}, a^*_{t+1}, s'_{t+1}) = 0$) according to (4.2). $V^*(s'_t) \geq V^*(s'_{t+1})$ follows immediately from the induction hypothesis and Lemma B.1.2, respectively. Then, we have $V^*(s_t) \geq Q(s_t, a_t = a^*_{t+1}) \geq V^*(s_{t+1})$. A straightforward inductive argument completes the proof. □

B.2 Numerical Results for Shorter EMS Response Times

In this appendix, we present the performance of our proposed tree-based selection strategy under shorter EMS response times and compare it with the results shown in Section 4.6. First, we provide an example of the tree-based selection strategy under a 6-minute EMS response in Section B.2.1. We then illustrate the performance of the tree-based strategy under shorter EMS response times compared to static strategies in Section B.2.2. Finally, we conclude this appendix with an analysis of the impact of a high acceptance rate on dispatch policies in Section B.2.3.

B.2.1 Tree-Based Strategy with Shorter EMS Response Times

Figure B.1 shows an example of the tree-based selection strategy trained on generated instances in Auckland with a 6-minute EMS response time and a 14% acceptance rate. This is a much shorter EMS response time, which SJAS aims to meet for 50% of incidents. Similar as Figure 4.3, the resulting tree-based strategy is highly interpretable. As we mentioned in Section 4.6.2, Figure B.1 is somewhat simpler, with fewer varieties of policies being selected compared to the tree-based strategy generated under a 12-minute EMS response time. *Keep n active until time t* policies are never selected because there is no time to wait for CFR response.

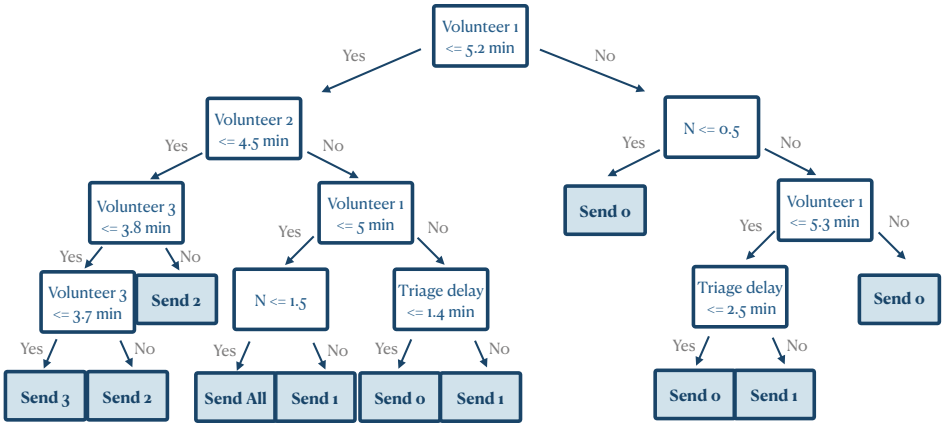


Figure B.1: Tree-based selection strategy for instances in Auckland with 14% acceptance rate and 6-minute EMS response time measured from call arrival.

When deciding which policy to use, Figure B.1 focuses more on the closer CFRs compared to Figure 4.3 because, with such a short response time after the system triage, only the closer CFRs could potentially contribute to patient survival. Additionally, CFR system triage delay now becomes more significant. In Figure 4.3, triage delay never becomes a feature that the tree splits on because CFRs still had a sufficient amount of time to respond, even after accounting for

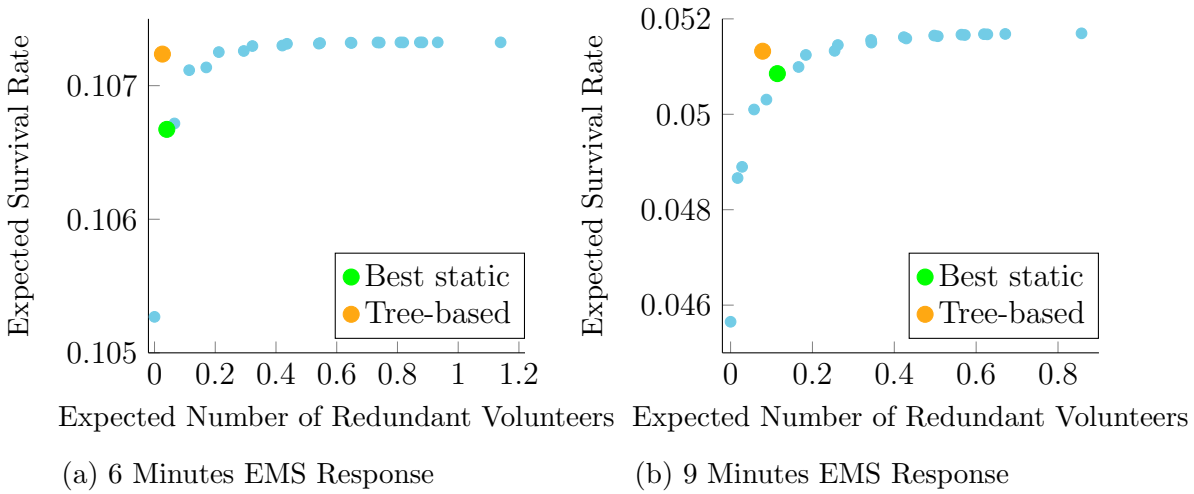
the triage delay. However, with a 6-minute EMS response time, triage delay can significantly impact the amount of time CFRs have left to respond, making it a more frequent feature in Figure B.1.

B.2.2 Adaptive Tree-Based Strategies vs. Static Strategies

Next, we compare the performance of our adaptive tree-based strategy with static strategies under both a 6-minute and a 9-minute EMS response times. We present an instance-wise plot showing the percentage improvement of the tree-selected policy over the best static policy under a 6-minute EMS response time. We compare these results to those shown in Section 4.6.2.

Figure B.2a and B.2b illustrate the trade-off between patient survival and redundant volunteers for both our adaptive and static strategies under a 14% acceptance rate, with 6-minute and 9-minute EMS response times, respectively. Under both EMS response times, the tree-based selection strategy lies far outside the efficient frontier of the static strategy, improving patient survival rates and reducing redundant volunteers. Notably, as shown in Figure B.2a, the tree-based selection strategy can achieve the survival rate of sending one more volunteer than the best static strategy, which is *Send 1 at time 0* under a 6-minute EMS response time, while having a smaller number of redundant volunteers.

Figure B.3 makes this observation more clear, showing the percentage improvement in patient survival versus the percentage improvement in the objective of the tree-selected policy compared to the best static policy for each test case under a 14% acceptance rate and a 6-minute EMS response time. Compared to Figure 4.5, the tree-based strategy under a 6-minute EMS response time improves upon the



(a) 6 Minutes EMS Response

(b) 9 Minutes EMS Response

Figure B.2: Trade-off between patient survival and number of redundant volunteers for static and adaptive tree-based strategies under a 14% acceptance rate. Each blue dot represents the performance of one static strategy. The orange dot indicates the performance of the tree-based strategy. The green dot indicates the performance of the best static strategy, which is *Send 1 at time 0* under a 6-minute EMS response time (left) and *Send 3 at time 0* under a 9-minute EMS response time (right).

best static strategy in most cases by increasing both the overall objective and patient survival through sending out more alerts. Only in a small number of cases, shown as the cluster of blue points in the lower left corner of Figure B.3, does the tree-selected policy choose to decrease the number of redundant volunteers. Given that the best static policy is *Send 1 at time 0*, this typically occurs when the time left for CFRs to respond is too short or when even the closest CFR is too far away. In these cases, it becomes optimal to rely solely on the ambulance response without any volunteer contribution. Compared to Figure 4.5, Figure B.3 shows that the tree-based strategy is most advantageous when there is a high CFR density within the dispatch radius. Since the best static policy is already pretty conservative in sending alerts, there is limited potential for improving overall objectives by reducing the number of alerts.

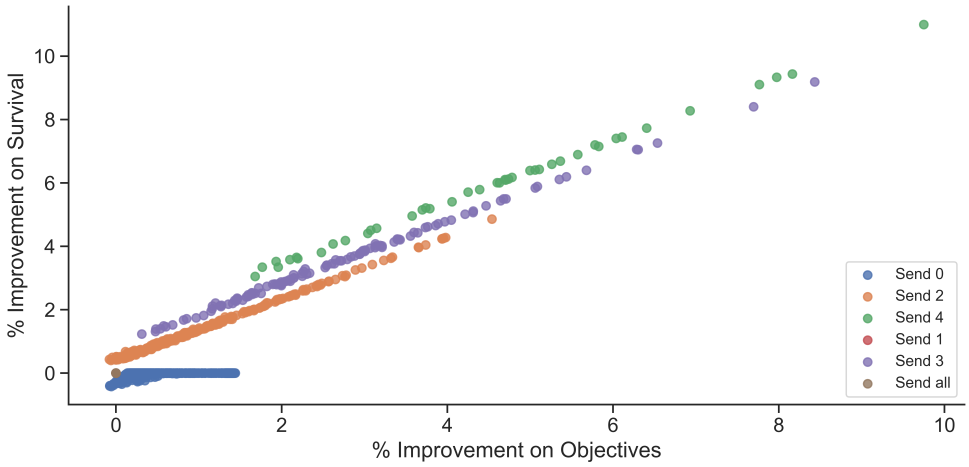
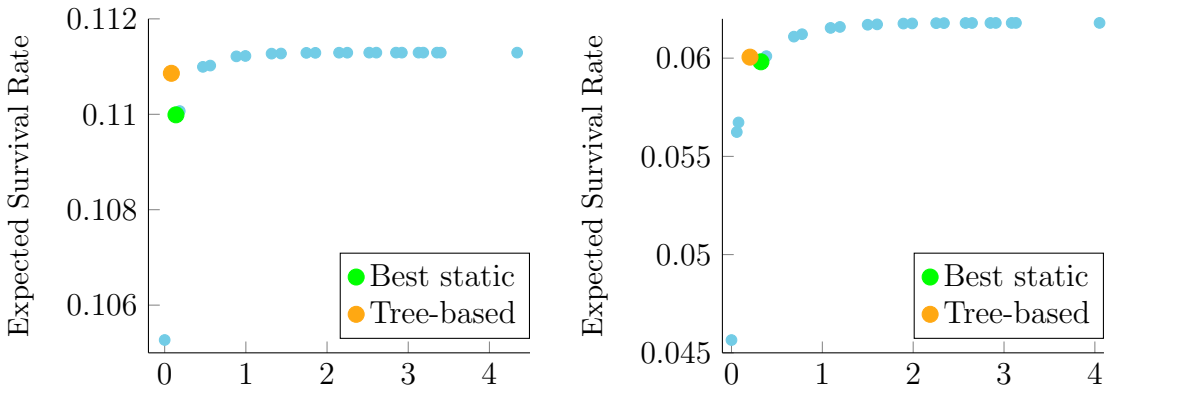


Figure B.3: Percentage improvement in survival rate vs. the percentage improvement in the objective of the tree-selected policy compared to the best static policy for 2000 test cases under a 14% acceptance and a 6-minute EMS response time.

B.2.3 Impact of High Volunteer Acceptance Rate

In the previous section, we assumed a 14% acceptance rate by directly bootstrapping CFR responses from the data. We now explore the impact of a higher volunteer acceptance rate on optimal volunteer dispatch strategies and their performance under shorter EMS response times. We assume a 50% acceptance rate, same as Section 4.6.3.

Figure B.4a and B.4b illustrate the trade-off between patient survival and redundant volunteers for both our adaptive and static strategies under a 50% acceptance rate, with 6-minute and 9-minute EMS response times, respectively. Unlike Figure 4.6, which improved upon the best static strategy by compromising survival rates, the tree-based selection strategies under 6-minute and 9-minute EMS response times are able to improve both patient survival rates and redundant volunteers. This is possibly due to the fact that with a shorter response time, the best static strategy has already refrained from sending out too many alerts, giving



(a) 6 Minutes EMS Response

(b) 9 Minutes EMS Response

Figure B.4: Trade-off between patient survival and number of redundant volunteers for static and adaptive tree-based strategies under a 50% acceptance rate. Each blue dot represents the performance of one static strategy. The orange dot indicates the performance of the selected tree-based strategy. The green dot indicates the performance of the best static strategy, which is *Send 2 at time 0* under a 6-minute EMS response time (right) and *Send 3 at time 0* under a 9-minute EMS response time (left).

the tree-based strategy less opportunity to improve solely by reducing redundant volunteers, even when high acceptance rates lead to more redundant volunteers.

Figure B.5 shows the percentage improvement in patient survival versus the percentage improvement in the objective of the tree-selected policy compared to the best static policy for each test case under a 50% acceptance rate and a 6-minute EMS response time. The best static strategy under both 14% and 50% acceptance rates with a 6-minute EMS response time is *Send 1 at time 0*. Similar to Figure B.3, and compared to Figure 4.7, the tree-based strategy under a 6-minute EMS response time improves upon the best static strategy in most cases by increasing patient survival, rather than saving on redundant volunteers. However, compared to Figure B.3, the cluster of blue points in the lower left corner of Figure B.5, where the tree-based strategy recognizes that a CFR response is unnecessary, shows a more significant improvement in the overall objective because the

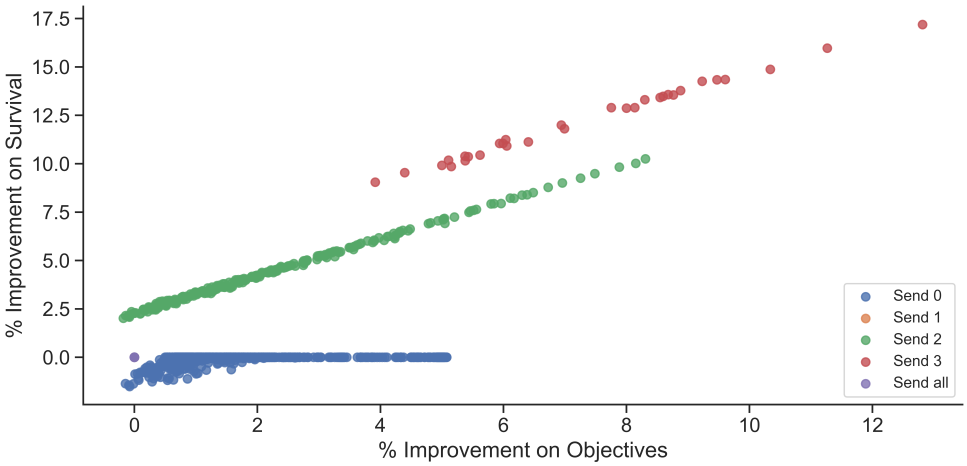


Figure B.5: Percentage improvement in survival rate vs. the percentage improvement in the objective of the tree-selected policy compared to the best static policy for 2000 test cases under a 50% acceptance and 6-minute EMS response time.

high acceptance rate leads to a higher probability of having redundant volunteers on-scene.

BIBLIOGRAPHY

Alban, A., Chick, S.E., Zoumpoulis, S.I., 2021. Expected value of information methods for contextual ranking and selection: Clinical trials and simulation optimization, in: 2021 Winter Simulation Conference (WSC), pp. 1–12. doi:10.1109/WSC52266.2021.9715303.

Ata, B., Lee, D., Sönmez, E., 2019a. Dynamic volunteer staffing in multicrop gleaning operations. Operations Research 67, 295–314.

Ata, B., Lee, D., Sönmez, E., 2019b. Dynamic volunteer staffing in multicrop gleaning operations. Operations Research 67, 295–314.

Bandara, D., Mayorga, M.E., McLay, L.A., 2012. Optimal dispatching strategies for emergency vehicles to increase patient survivability. International Journal of Operational Research 15, 195–214.

Barbour, A.D., Holst, L., Janson, S., 1992. Poisson approximation. volume 2 of *Oxford studies in probability*. Oxford University Press.

Berahas, A.S., Cao, L., Scheinberg, K., 2021. Global convergence rate analysis of a generic line search algorithm with noise. SIAM Journal on Optimization 31, 1489–1518. URL: https://doi.org/10.1137/19M1291832, doi:10.1137/19M1291832, arXiv:https://doi.org/10.1137/19M1291832.

Berdowski, J., Berg, R.A., Tijssen, J.G., Koster, R.W., 2010. Global incidences of out-of-hospital cardiac arrest and survival rates: Systematic review of 67 prospective studies. Resuscitation 81, 1479–1487.

van den Berg, P.L., Henderson, S.G., Jagtenberg, C.J., Li, H., 2024a. Modeling the impact of community first responders. Management Science .

van den Berg, P.L., Henderson, S.G., Li, H., Dicker, B., Jagtenberg, C.J., 2024b. Community first response for cardiac arrest: comparing phased dispatch policies through monte carlo simulation. URL: https://www.medrxiv.org/content/10.1101/2024.01.17.24301457v1. accessed: 2024-05-14.

Bertsimas, D., Dunn, J., . Optimal classification trees 106, 1039–1082. URL: https://doi.org/10.1007/s10994-017-5633-9, doi:10.1007/s10994-017-5633-9.

Bomze, I.M., Rinaldi, F., Zeffiro, D., 2020. Active set complexity of the away-step Frank–Wolfe algorithm. SIAM Journal on Optimization 30, 2470–2500. URL: https://doi.org/10.1137/19M1309419, doi:10.1137/19M1309419.

Boutilier, J.J., Brooks, S.C., Janmohamed, A., Byers, A., Buick, J.E., Zhan, C., Schoellig, A.P., Cheskes, S., Morrison, L.J., Chan, T.C.Y., 2017. Optimizing a drone network to deliver automated external defibrillators. Circulation 135, 2454–2465.

Boutilier, J.J., Chan, T.C.Y., 2022. Drone network design for cardiac arrest response. Manufacturing & Service Operations Management 24, 2407–2424. doi:10.1287/msom.2022.1092.

Brandweer Amsterdam-Amstelland, 2021. Beleidsplan 2021-2024. https://openresearch.amsterdam/image/2021/7/7/eindconcept_vraa_beleidsplan_brandweer.pdf.

Bray, J., Howell, S., Ball, S., Doan, T., Bosley, E., Smith, K., Dicker, B., Faddy, S., Thorrowgood, M., Swain, A., et al., 2022. The epidemiology of out-of-hospital cardiac arrest in Australia and New Zealand: A binational report from the Australasian Resuscitation Outcomes Consortium (Aus-ROC). Resuscitation 172, 74–83.

Brooks, S.C., Simmons, G., Worthington, H., Bobrow, B.J., Morrison, L.J., 2016. The PulsePoint Respond mobile device application to crowdsource basic life support for patients with out-of-hospital cardiac arrest: challenges for optimal implementation. Resuscitation 98, 20–26.

Buter, R., Koffijberg, E., Stieglis, R., Demirtas, D., 2024. Health economic evaluation of data-driven aed locations. Under Review .

Cakmak, S., Zhou, E., Gao, S., 2021. Contextual ranking and selection with gaussian processes, in: 2021 Winter Simulation Conference (WSC), pp. 1–12. doi:10.1109/WSC52266.2021.9715499.

Chan, T.C., Demirtas, D., Kwon, R.H., 2016. Optimizing the deployment of public

access defibrillators. Management Science 62, 3617–3635.

Chan, T.C., Shen, Z.J.M., Siddiq, A., 2018. Robust defibrillator deployment under cardiac arrest location uncertainty via row-and-column generation. Operations Research 66, 358–379.

Chu, J., Leung, K.B., Snobelen, P., Nevils, G., Drennan, I.R., Cheskes, S., Chan, T.C., 2021. Machine learning-based dispatch of drone-delivered defibrillators for out-of-hospital cardiac arrest. Resuscitation 162, 120–127.

Cont, R., Kotlicki, A., Xu, R., 2021. Modelling COVID-19 contagion: risk assessment and targeted mitigation policies. Royal Society Open Science 8, 201535.

Daley, D., Vere-Jones, D., 2008. An Introduction to the Theory of Point Processes: Volume II: General Theory and Structure. Probability and Its Applications, Springer New York.

Daskin, M., 1983. A maximum expected location model: Formulation, properties and heuristic solution. Transportation Science 7, 48–70.

Dicker, B., Garrett, N., Wong, S., McKenzie, H., McCarthy, J., Jenkin, G., Smith, T., Skinner, J.R., Pegg, T., Devlin, G., et al., 2019. Relationship between socioeconomic factors, distribution of public access defibrillators and incidence of out-of-hospital cardiac arrest. Resuscitation 138, 53–58.

Du, J., Gao, S., Chen, C.H., 2023. A contextual ranking and selection method for personalized medicine. Manufacturing & Service Operations Management 26, 167–181. URL: https://doi.org/10.1287/msom.2022.0232, doi:10.1287/msom.2022.0232.

Erkut, E., Ingolfsson, A., Erdoğan, G., 2008. Ambulance location for maximum survival. Naval Research Logistics 55, 42–58.

Escallon-Barrios, M., Noham, R., Smilowitz, K., 2024. Dual mode scheduling in volunteer management. Socio-Economic Planning Sciences 92, 101796.

Facebook Data For Good, 2021. Colocation maps. https://dataforgood.facebook.com/dfg/tools/colocation-maps.

Folke, F., Lippert, F.K., Nielsen, S.L., Gislason, G.H., Hansen, M.L., Schramm, T.K., Sørensen, R., Fosbøl, E.L., Andersen, S.S., Rasmussen, S., et al., 2009. Clinical perspective. Circulation 120, 510–517.

Frank, M., Wolfe, P., 1956. An algorithm for quadratic programming. Naval Research Logistics Quarterly 3, 95–110.

Global Resuscitation Alliance, 2019. GoodSAM - a smart phone application to crowd-source CPR in New Zealand. https://www.globalresuscitationalliance.org/wp-content/uploads/2019/12/St_John_NZ_Smart_Tech.pdf.

GoodSAM Platform, 2020. GoodSAM. https://www.goodsamapp.org/. [Online; accessed 29-02-2020].

van der Hagen, L., Agatz, N., Spliet, R., Visser, T.R., Kok, L., 2022. Machine learning–based feasibility checks for dynamic time slot management. Transportation Science 58, 94–109. URL: https://doi.org/10.1287/trsc.2022.1183, doi:10.1287/trsc.2022.1183.

HartslagNu, 2020. Hartslagnu. https://hartslagnu.nl/. [Online; accessed 14-12-2020].

Haskins, B., Nehme, Z., Dicker, B., Wilson, M.H., Ray, M., Bernard, S., Cameron, P., Smith, K., 2021. A binational survey of smartphone activated volunteer responders for out-of-hospital cardiac arrest: availability, interventions, and post-traumatic stress. Resuscitation 169, 67–75.

Henderson, S.G., Berg, P.L., Jagtenberg, C.J., Li, H., 2022. How should volunteers be dispatched to out-of-hospital cardiac arrest cases? Queueing Systems , 1–3.

Jansma, T., 2014. Extending the emergency medical services network for out-of-hospital cardiac arrest victims: an explorative study for the province of Drenthe. Master's thesis. University of Groningen.

Johari, R., Kamble, V., Kanoria, Y., 2021. Matching while learning. Operations Research 69, 655–681. URL: https://doi.org/10.1287/opre.2020.2013, doi:10.1287/opre.2020.2013, arXiv:https://doi.org/10.1287/opre.2020.2013.

Kazemi, L., Shahabi, C., 2012. Geocrowd: enabling query answering with spatial

crowdsourcing, in: Proceedings of the 20th International Conference on Advances in Geographic Information Systems, Association for Computing Machinery, New York, NY, USA. p. 189–198. URL: https://doi.org/10.1145/2424321.2424346, doi:10.1145/2424321.2424346.

Keslin, G., Nelson, B.L., Plumlee, M., Pagnoncelli, B.K., Rahimian, H., 2022. A classification method for ranking and selection with covariates, in: 2022 Winter Simulation Conference (WSC), IEEE. pp. 1–12.

Khalemsky, M., Schwartz, D.G., 2017. Emergency response community effectiveness: a simulation modeler for comparing emergency medical services with smartphone-based samaritan response. Decision Support Systems 102, 57–68.

Kingman, J.F.C., 1993. Poisson processes. Oxford University Press.

Kragholm, K., Wissenberg, M., Mortensen, R.N., Hansen, S.M., Malta Hansen, C., Thorsteinsson, K., Rajan, S., Lippert, F., Folke, F., Gislason, G., et al., 2017. Bystander efforts and 1-year outcomes in out-of-hospital cardiac arrest. New England Journal of Medicine 376, 1737–1747.

Lancaster, G., Herrmann, J.W., 2021. Computer simulation of the effectiveness of novel cardiac arrest response systems. Resuscitation Plus 7, 100153.

Larson, R.C., Odoni, A.R., 1981. Urban operations research. Prentice Hall, Englewood Cliffs NJ.

Liu, W., Sun, Q., Ching Tang, L., Ye, Z., 2022. Robust data-driven design of a smart cardiac arrest response system. under review at Production and Operations Management .

Lives, n.d. Emergency responders. https://www.lives.org.uk/what-we-do/emergency-responders/. [Online; accessed 24-10-2022].

De Maio, V., Stiell, I., Wells, G., Spaite, D., 2003. Optimal defibrillation for maximum out-of-hospital cardiac arrest survival rates. Annals of Emergency Medicine 42, 242–250.

Manshadi, V., Rodilitz, S., 2020. Online policies for efficient volunteer crowdsourcing,

in: Proceedings of the 21st ACM Conference on Economics and Computation, pp. 315–316.

Manshadi, V., Rodilitz, S., 2022. Online policies for efficient volunteer crowdsourcing. Management Science 68, 6572–6590. URL: https://doi.org/10.1287/mnsc.2021.4220, doi:10.1287/mnsc.2021.4220, arXiv:https://doi.org/10.1287/mnsc.2021.4220.

Matinrad, N., Granberg, T., Ennab Vogel, N., Angelakis, V., 2019. Optimal dispatch of volunteers to out-of-hospital cardiac arrest patients, in: Proceedings of the 52nd Hawaii International Conference on System Sciences.

Matinrad, N., Granberg, T.A., 2023. Optimal pre-dispatch task assignment of volunteers in daily emergency response. Socio-Economic Planning Sciences 87, 101589.

Matinrad, N., Granberg, T.A., Angelakis, V., 2021a. Modeling uncertain task compliance in dispatch of volunteers to out-of-hospital cardiac arrest patients. Computers & Industrial Engineering 159, 107515.

Matinrad, N., Granberg, T.A., Angelakis, V., 2021b. Modeling uncertain task compliance in dispatch of volunteers to out-of-hospital cardiac arrest patients. Computers & Industrial Engineering 159, 107515.

McElfresh, D.C., Kroer, C., Pupyrev, S., Sodomka, E., Sankararaman, K.A., Chauvin, Z., Dexter, N., Dickerson, J.P., 2020. Matching algorithms for blood donation, in: Proceedings of the 21st ACM Conference on Economics and Computation, pp. 463–464.

McLay, L.A., 2009. Emergency medical service systems that improve patient survivability, in: Wiley Encyclopedia of Operations Research and Management Science. Wiley.

Ministry of Health, 2020. Emergency ambulance service national performance report. https://www.health.govt.nz/system/files/documents/pages/naso-march-report-2019-2020.pdf.

Nazarian, A., 2018. Optimizing the deployment of automated external defibrillators by

a data-driven algorithmic approach. Master's thesis. University of Twente.

NHS North West Ambulance Service, n.d. Community first responders (cfr). https://www.nwas.nhs.uk/get-involved/volunteering/community-first-responder/. [Online; accessed 24-10-2022].

Nichol, G., Stiell, I.G., Laupacis, A., De Maio, V.J., Wells, G.A., et al., 1999. A cumulative meta-analysis of the effectiveness of defibrillator-capable emergency medical services for victims of out-of-hospital cardiac arrest. Annals of Emergency Medicine 34, 517–525.

Nisingizwe, M.P., Ndishimye, P., Swaibu, K., Nshimiyimana, L., Karame, P., Dushimiyimana, V., Musabyimana, J.P., Musanabaganwa, C., Nsanzimana, S., Law, M.R., 2022. Effect of unmanned aerial vehicle (drone) delivery on blood product delivery time and wastage in Rwanda: a retrospective, cross-sectional study and time series analysis. The Lancet Global Health 10, 564–569. doi:10.1016/S2214-109X(22)00048-1.

Northern Ireland Ambulance Service Health & Social Care Trust, n.d. Community first responder scheme. https://nias.hscni.net/our-services/involving-our-community/community-first-responder-scheme/. [Online; accessed 24-10-2022].

Oving, I., Masterson, S., Tjelmeland, I.B., Jonsson, M., Semeraro, F., Ringh, M., Truhlar, A., Cimpoesu, D., Folke, F., Beesems, S.G., et al., 2019. First-response treatment after out-of-hospital cardiac arrest: a survey of current practices across 29 countries in Europe. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine 27, 112.

Özkan, E., Ward, A.R., 2020. Dynamic matching for real-time ride sharing. Stochastic Systems 10, 29–70. URL: https://doi.org/10.1287/stsy.2019.0037, doi:10.1287/stsy.2019.0037, arXiv:https://doi.org/10.1287/stsy.2019.0037.

Phung, V.H., Trueman, I., Togher, F., Orner, R., Siriwardena, A.N., 2017. Community first responders and responder schemes in the United Kingdom: systematic scoping

review. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine 25, 1–9.

Pijls, R.W., Nelemans, P.J., Rahel, B.M., Gorgels, A.P., 2016. A text message alert system for trained volunteers improves out-of-hospital cardiac arrest survival. Resuscitation 105, 182–187.

Pijls, R.W., Nelemans, P.J., Rahel, B.M., Gorgels, A.P., 2019. Characteristics of a novel citizen rescue system for out-of-hospital cardiac arrest in the Dutch province of Limburg: relation to incidence and survival. Netherlands Heart Journal 27, 100–107.

PulsePoint, 2020. Pulsepoint. https://www.pulsepoint.org/. [Online; accessed 29-02-2020].

Qin, Z., Tang, X., Jiao, Y., Zhang, F., Xu, Z., Zhu, H., Ye, J., 2020. Ride-hailing order dispatching at didi via reinforcement learning. INFORMS Journal on Applied Analytics 50, 272–286.

Radiolab, 2024. Radiolab. https://radiolab.org/podcast/how-to-save-a-life. [Online; accessed 29-07-2024].

Resnick, S.I., 2002. Adventures in Stochastic Processes. Springer New York.

Ridler, S., 2020. Emergency Medical Services Research at the University of Auckland, Department of Engineering Science. https://github.com/uoa-ems-research.

Ridler, S., Mason, A.J., Raith, A., 2017. A simulation package for emergency medical services, in: Proceedings of the 51st Annual Conference of the ORSNZ.

Ridler, S., Mason, A.J., Raith, A., 2022. A simulation and optimisation package for emergency medical services. European Journal of Operational Research 298, 1101–1113.

Ringh, M., Fredman, D., Nordberg, P., Stark, T., Hollenberg, J., 2011. Mobile phone technology identifies and recruits trained citizens to perform CPR on out-of-hospital cardiac arrest victims prior to ambulance arrival. Resuscitation 82, 1514–1518.

Safegraph, 2021. Safegraph. https://www.safegraph.com/.

Sampson, S.E., 2006. Optimization of volunteer labor assignments. Journal of Operations Management 24, 363–377. URL: https://onlinelibrary.wiley.com/doi/abs/10.1016/j.jom.2005.05.005, doi:https://doi.org/10.1016/j.jom.2005.05.005, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1016/j.jom.2005.05.005.

Sasson, C., Rogers, M.A., Dahl, J., Kellermann, A.L., 2010. Predictors of survival from out-of-hospital cardiac arrest: a systematic review and meta-analysis. Circulation: Cardiovascular Quality and Outcomes 3, 63–81.

Serfling, R.J., 1980. Approximation Theorems of Mathematical Statistics. Wiley Series in Probability and Mathematical Statistics, Wiley & Sons, New York.

Shin, J., Marla, L., Boutilier, J.J., 2022. Heterogeneous facility location under coordination and collaboration: The case of ambulance-bystander-drone coordination. under review .

Skogvoll, E., Lindqvist, B.H., 1999. Modeling the occurrence of cardiac arrest as a poisson process. Annals of emergency medicine 33, 409–417.

Slaa, G., 2020. Increasing cardiac arrest survival by improving the volunteer alerting algorithm. URL: http://essay.utwente.nl/80432/.

Slaa, G., 2020. Increasing cardiac arrest survival by improving the volunteer alerting algorithm. Master's thesis. University of Twente.

Smith, C.M., Wilson, M.H., Ghorbangholi, A., Hartley-Sharpe, C., Gwinnutt, C., Dicker, B., Perkins, G.D., 2017. The use of trained volunteers in the response to out-of-hospital cardiac arrest–the GoodSAM experience. Resuscitation 121, 123–126.

Stats New Zealand Geographic Data Service, 2015a. Age by meshblock (2013 census). https://datafinder.stats.govt.nz/layer/8447-age-by-meshblock-2013-census/data/. [Online; accessed 25-10-2021].

Stats New Zealand Geographic Data Service, 2015b. Population by meshblock 2013. https://datafinder.stats.govt.nz/layer/8437-population-by-meshblock-2013-census/. [Online; accessed 28-10-2020].

Stats New Zealand Geographic Data Service, 2016. Area units 2013. https://datafinder.stats.govt.nz/layer/25743-area-unit-2013/. [Online; accessed 28-10-2020].

Stats New Zealand Geographic Data Service, 2020. 2018 census main means of travel to work by statistical area 2. https://datafinder.stats.govt.nz/table/104720-2018-census-main-means-of-travel-to-work-by-statistical-area-2/. [Online; accessed 25-10-2021].

Stieglis, R., Zijlstra, J.A., Riedijk, F., Smeekes, M., van der Worp, W.E., Koster, R.W., 2020. AED and text message responders density in residential areas for rapid response in out-of-hospital cardiac arrest. Resuscitation 150, 170–177.

Tafreshian, A., Masoud, N., Yin, Y., 2020. Frontiers in service science: Ride matching for peer-to-peer ride sharing: A review and future directions. Service Science 12, 44–60.

Tong, Y., Zhou, Z., Zeng, Y., Chen, L., Shahabi, C., . Spatial crowdsourcing: a survey 29, 217–250. URL: https://doi.org/10.1007/s00778-019-00568-7, doi:10.1007/s00778-019-00568-7.

Valenzuela, T.D., Roe, D.J., Cretin, S., Spaite, D.W., Larsen, M.P., 1997. Estimating effectiveness of cardiac arrest interventions: a logistic regression survival model. Circulation 96, 3308–3313.

Waalewijn, R.A., de Vos, R., Tijssen, J.G., Koster, R.W., 2001. Survival models for out-of-hospital cardiopulmonary resuscitation from the perspectives of the bystander, the first responder, and the paramedic. Resuscitation 51, 113–122.

Wikipedia, 2020. Burgerhulpverlening. https://nl.wikipedia.org/wiki/Burgerhulpverlening. [Online; accessed 14-12-2020].

Yan, S., Gan, Y., Jiang, N., Wang, R., Chen, Y., Luo, Z., Zong, Q., Chen, S., Lv, C., 2020. The global survival rate among adult out-of-hospital cardiac arrest patients who received cardiopulmonary resuscitation: a systematic review and meta-analysis. Critical Care 24.

Zaffar, M.A., Rajagopalan, H.K., Saydam, C., Mayorga, M., Sharer, E., 2016. Coverage, survivability or response time: A comparative study of performance statistics used in ambulance location models via simulation optimization. Operations Research for Health Care 11, 1–12.

Zijlstra, J.A., Stieglis, R., Riedijk, F., Smeekes, M., Van der Worp, W.E., Koster, R.W., 2014. Local lay rescuers with AEDs, alerted by text messages, contribute to early defibrillation in a Dutch out-of-hospital cardiac arrest dispatch system. Resuscitation 85, 1444–1449.