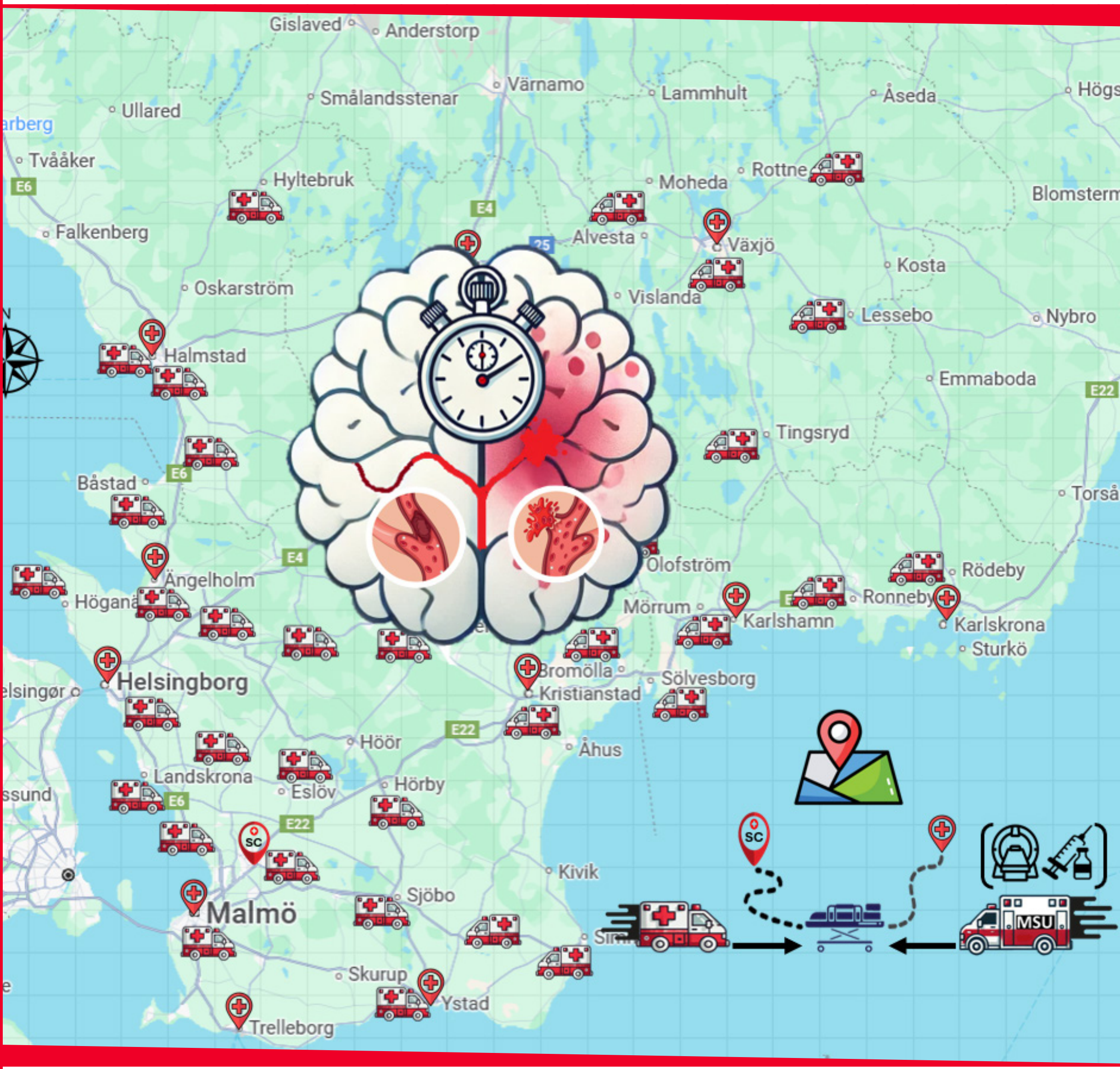
STUDIES IN COMPUTER SCIENCE 35

MALMÖ UNIVERSITY



Saeid Amouzad Mahdiraji

Optimization and Simulation Modeling for Improved Analysis and Planning of Prehospital Stroke Care

This thesis explores how simulation and optimization modeling can enhance decision-making and planning in prehospital stroke care. With a focus on Mobile Stroke Units (MSUs)—specialized ambulances equipped for on-site stroke diagnosis and treatment—it introduces new approaches to address logistical challenges and reduce treatment delays. The research presents optimization models for MSU placement, develops a simulation framework for evaluating stroke transport policies, and integrates machine learning for dynamic ambulance travel time estimation. Applied to Sweden's Southern Healthcare Region, the findings provide valuable insights for improving emergency medical services and patient outcomes in stroke care.

MALMÖ UNIVERSITY PRESS
205 06 MALMÖ, SWEDEN
MAU.SE

ISBN 978-91-7877-596-5
DOI 10.24834/isbn.9789178775972

Printed by Media-Tryck, Lund 2025 NORDIC SWAN ECOLABEL 3041 0903

OPTIMIZATION AND SIMULATION MODELING FOR IMPROVED ANALYSIS AND PLANNING OF PREHOSPITAL STROKE CARE

Optimization and Simulation Modeling for Improved Analysis and Planning of Prehospital Stroke Care

Thesis for Doctoral Degree (PhD)

By Saeid Amouzad Mahdiraji

The thesis will be defended in public in NIC0319,
Niagara, Malmö University
On Thursday, March 27, 2025, at 14:00

Principal Supervisor
Associate Professor Johan Holmgren
Malmö University
Department of Computer Science and Media Technology

Co-Supervisor
Associate Professor Radu-Casian Mihailescu
Malmö University | Heriot-Watt University
Department of Computer Science and Media Technology | School of Mathematical and Computer Sciences

Co-Supervisor
Professor Jesper Petersson
Lund University | Skåne University Hospital
Department of Neurology | Department of Healthcare Management

Opponent
Associate Professor Derya Demirtas
University of Twente, Netherlands
Department of Industrial Engineering and Business Information Systems

Examination Board
Professor Tobias Andersson Granberg
Linköping University, Sweden
Department of Engineering and Natural Sciences

Professor Sebastiaan Meijer
KTH Royal Institute of Technology, Sweden
Department of Biomedical Engineering and Health Systems

Professor Holger Billhardt
Rey Juan Carlos University, Spain
Department of Computer Science and Statistics

Malmö University
Studies in Computer Science No 35,
PhD Dissertation 2025

Cover design by the author

DOI 10.24834/isbn.9789178775972
ISBN 978-91-7877-596-5 (print)
ISBN 978-91-7877-597-2 (pdf)

Printed in Sweden by Media-Tryck, Lund University,
Lund 2025

Media-Tryck is a Nordic Swan Ecolabel certified provider of printed material. Read more about our environmental work at www.mediatryck.lu.se

MADE IN SWEDEN

SAEID AMOUZAD MAHDIRAJI

OPTIMIZATION AND SIMULATION MODELING FOR IMPROVED ANALYSIS AND PLANNING OF PREHOSPITAL STROKE CARE

Malmö University
Faculty of Technology and Society
Department of Computer Science and Media Technology

Studies in Computer Science
Faculty of Technology and Society
Malmö University

1. Jevinger, Åse. *Toward intelligent goods: characteristics, architectures and applications*, 2014, Doctoral dissertation.
2. Dahlskog, Steve. *Patterns and procedural content generation in digital games: automatic level generation for digital games using game design patterns*, 2016, Doctoral dissertation.
3. Fabijan, Aleksander. *Developing the right features: the role and impact of customer and product data in software product development*, 2016, Licentiate thesis.
4. Paraschakis, Dimitris. *Algorithmic and ethical aspects of recommender systems in e-commerce*, 2018, Licentiate thesis.
5. Hajinasab, Banafsheh. *A Dynamic Approach to Multi Agent Based Simulation in Urban Transportation Planning*, 2018, Doctoral dissertation.
6. Fabijan, Aleksander. *Data-Driven Software Development at Large Scale*, 2018, Doctoral dissertation.
7. Bugeja, Joseph. *Smart Connected Homes: Concepts, Risks, and Challenges*, 2018, Licentiate thesis.
8. Alkhabbas, Fahed. *Towards Emergent Configurations in the Internet of Things*, 2018, Licentiate thesis.
9. Paraschakis, Dimitris. *Sociotechnical Aspects of Automated Recommendations: Algorithms, Ethics, and Evaluation*, 2020, Doctoral dissertation.
10. Tegen, Agnes. *Approaches to Interactive Online Machine Learning*, 2020, Licentiate thesis.
11. Alvarez, Alberto. *Exploring the Dynamic Properties of Interaction in Mixed-Initiative Procedural Content Generation*, 2020, Licentiate thesis.
12. Alkhabbas, Fahed. *Realizing Emergent Configurations in the Internet of Things*, 2020, Doctoral dissertation.
13. Ashouri, Majid. *Towards Supporting IoT System Designers in Edge Computing Deployment Decisions*, 2021, Licentiate thesis.
14. Bugeja, Joseph. *On Privacy and Security in Smart Connected Homes*, 2021, Doctoral dissertation.
15. Azadvar, Ahmad. *Predictive Psychological Player Profiling*, 2021, Licentiate thesis.
16. Holmberg, Lars. *Human in Command Machine Learning*, 2021, Licentiate thesis.
17. John, Meenu Mary. *Design Methods and Processes for ML/DL Models*, 2021, Licentiate thesis.
18. Tegen, Agnes. *Interactive Online Machine Learning*, 2022, Doctoral dissertation.
19. Ouhaichi, Hamza. *Towards Designing a Flexible Multimodal Learning Analytics System*, 2022, Licentiate thesis.
20. Alvarez, Alberto. *Exploring Game Design through Human-AI Collaboration*, 2022, Doctoral dissertation.
21. Amouzad Mahdiraji, Saeid. *On the Use of Simulation and Optimization for the Analysis and Planning of Prehospital Stroke Care*, 2022, Licentiate thesis.
22. Holmberg, Lars. *Neural Networks in Context: Challenges and Opportunities*, 2023, Doctoral dissertation.
23. Pettersson, Mårten. *Getting Engaged in Cooperation: Design, Distance and Distributed Work*, 2023, Doctoral dissertation.
24. Engström, Jimmy. *Scaling Indoor Positioning - Methods for improving accuracy and privacy of indoor positioning,* 2023, Licentiate thesis.
25. Dytckov, Sergei. *Modelling and Simulating Demand-Responsive Transport*, 2023, Licentiate thesis.
26. Ouhaichi, Hamza. *A Framework for Designing and Analysing Multimodal Learning Analytics Systems*, 2024, Doctoral dissertation.
27. Ymeri, Gent. *Machine Learning-Driven Analysis of Sensor Data for Objective Assessment of Parkinson's Disease Motor Symptoms in Home Environments*, 2024, Licentiate thesis.
28. Brötzner, Anna. *Visibility in Graphs and Polygons*, 2024, Licentiate thesis.

29. Bagheri, Sally. *The Role of Ethics in Smart Homes*, 2024, Licentiate thesis.
30. John, Meenu Mary. *Towards Continuous Development of MLOps Practices*, 2025, Doctoral dissertation.
31. Belfrage, Michael. *Agent-based Social Simulation & Policy-Modelling: Facilitating Realistic and Credible Decision-making Support*, 2025, Licentiate thesis.
32. Johansson, Emil. *Towards More Credible Agent-based Models for Policy Support*, 2025, Licentiate thesis.
33. Abid, Muhammad Adil. *Artificial Intelligence for Enhanced Prehospital Stroke Care*, 2025, Licentiate thesis.
34. Saleem, Hajira. *Towards Improving Localization for Autonomous Vehicles*, 2025, Licentiate thesis.
35. Amouzad Mahdiraji, Saeid. *Optimization and Simulation Modeling for Improved Analysis and Planning of Prehospital Stroke Care*, 2025, Doctoral dissertation.

This publication also available at:
mau.diva-portal.org

To Hanna, my wife, and my parents

TABLE OF CONTENTS

ABSTRACT

Rapid treatment is crucial for minimizing the consequences of a stroke. However, logistical challenges and the complexity of accurate stroke diagnosis often impede timely and effective treatment. One way to reduce time to treatment is the use of so-called mobile stroke units (MSUs), which are specialized ambulances equipped to diagnose and treat stroke patients on site. The adequate planning and optimization of prehospital stroke transport policies involving MSUs can help reduce delays in accessing treatment. Mathematical optimization and simulation are useful approaches for optimizing and assessing different stroke transport policies without endangering patient's health.

The aim of this thesis is to explore how optimization and simulation can improve the analysis and planning of prehospital stroke care. Specifically, optimization is used to determine optimal MSU placements, while simulation is applied to evaluate stroke transport policies, including those involving MSUs. To achieve this aim, the thesis is structured around four main objectives, in which we develop and analyze a number of different optimization and simulation models. First, the MSU placement problem is solved using an exhaustive search algorithm and formulated as a mixed-integer linear programming model to determine optimal MSU placements. The objective of solving this problem is to make a trade-off between efficiency and equity, ensuring maximum population coverage and equitable service across a region. Second, macro-level and micro-level simulation models are proposed to evaluate various stroke transport policies, including MSUs. Third, a simulation modeling framework is introduced to enable the construction of discrete event simulation models for emergency medical services (EMS) policy analysis, supporting flexible and adaptive simulations of real-world EMS operations. The framework incorporates various decision policies, such as emergency vehicle selection, dispatch type (single and co-dispatch) selection, and hospital selection, allowing for the evaluation of

stroke transport policies across different stroke types. Lastly, dynamic travel time calculations and machine learning-based travel time estimations are integrated into the framework to enhance the flexibility and reliability of EMS simulations.

Through scenario studies conducted in Sweden's Southern Healthcare Region, this research demonstrates how optimization and simulation can support effective stroke transport policy planning and improve decision-making in prehospital stroke care. The identified MSU placements, along with the evaluated dispatch policies, highlight significant potential for reducing the time to diagnosis and treatment for different types of strokes. Faster time to treatment not only enhances overall stroke care delivery but also improves patient outcomes by reducing stroke-related disabilities. The findings underscore the value of these approaches in guiding EMS policy design, ultimately contributing to better patient outcomes and reduced social impacts of stroke. The results of this thesis aim to assist public health authorities in making informed decisions to optimize prehospital stroke care.

Keywords: Stroke Transport Policy, Mobile Stroke Unit, MSU, Optimization, Simulation, Prehospital Stroke Care, Modeling Framework, Emergency Medical Services, Dynamic Travel Time, Machine Learning, Ambulance Travel Time Estimation

LIST OF PUBLICATIONS

Included Papers:

Paper I. **S. Amouzad Mahdiraji**, O. Dahllöf, F. Hofwimmer, J. Holmgren, R.-C. Mihailescu, and J. Petersson, "Mobile Stroke Units for Acute Stroke Care in the South of Sweden," *Cogent Engineering*, vol. 8, no. 1, 2021, doi: https://doi.org/10.1080/23311916.2021.1874084.

Paper II. **S. Amouzad Mahdiraji**, J. Holmgren, R.-C. Mihailescu, and J. Petersson, "An Optimization Model for the Trade-off Between Efficiency and Equity for Mobile Stroke Unit Placement," in *9th KES International Conference on Innovation in Medicine and Healthcare (KES-InMed-21)*, pp. 183-193. Springer, Singapore, 2021, doi: https://doi.org/10.1007/978-981-16-3013-2_15.

Paper III. **S. Amouzad Mahdiraji**, J. Holmgren, R.-C. Mihailescu, and J. Petersson, "A Micro-Level Simulation Model for Analyzing the Use of MSUs in Southern Sweden," in *Proceedings of the 11th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH 2021)*, vol. 198, pp. 132-139, 2022, doi: https://doi.org/10.1016/j.procs.2021.12.220.

Paper IV. **S. Amouzad Mahdiraji**, J. Holmgren, A. Alshaban, R.-C. Mihailescu, J. Petersson, and J. Al Fatah, "A Framework for Constructing Discrete Event Simulation Models for Emergency Medical Service Policy Analysis," in *Proceedings of the 12th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH 2022)*, vol. 210, pp. 133-140, 2022, doi: https://doi.org/10.1016/j.procs.2022.10.129.

Paper V. **S. Amouzad Mahdiraji**, M. A. Abid, J. Holmgren, R.-C. Mihailescu, F. Lorig, and J. Petersson, "An Optimization Model for the Placement of Mobile Stroke Units," in *International Conference on Advanced Research in Technologies, Information, Innovation and Sustainability*, pp. 297-310. Springer Nature, Switzerland, 2023, doi: https://doi.org/10.1007/978-3-031-48858-0_24.

Paper VI. **S. Amouzad Mahdiraji**, J. Holmgren, R.-C. Mihailescu, and J. Petersson, "Simulation-based Analysis of Co-dispatching in Prehospital Stroke Care," in *Proceedings of the 15th International Conference on Ambient Systems, Networks and Technologies (ANT 2024)*, vol. 238, pp. 412-419, 2024, doi: https://doi.org/10.1016/j.procs.2024.06.042.

Paper VII. **S. Amouzad Mahdiraji**, M. Juninger, N. Narvell, J. Holmgren, R.-C. Mihailescu, and J. Petersson, "Implementing Dynamic Travel Time Calculation in EMS Simulations: Impacts on Prehospital Stroke Care and Transportation," accepted for publication in *proceedings of International Conference on Health and Social Care Information Systems and Technologies (HCist 2024)*, 2024.

Paper VIII. **S. Amouzad Mahdiraji**, M. A. Abid, and J. Holmgren, "Integrating Machine Learning-Based Ambulance Travel Time Estimation into an Emergency Medical Services Simulation Modeling Framework," in *Proceedings of the 14th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH 2024)*, vol. 251, pp. 479-486, 2024, doi: https://doi.org/10.1016/j.procs.2024.11.136.

The published papers have been reformatted to fit the layout of this thesis.

Related papers not included in the thesis

Paper IX. M. A. Abid, **S. Amouzad Mahdiraji**, F. Lorig, J. Holmgren, R.-C. Mihailescu, and J. Petersson, “A Genetic Algorithm for Optimizing Mobile Stroke Unit Deployment,” in *Proceedings of the 27th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems*, vol. 225, pp. 3536-3545, 2023, doi: https://doi.org/10.1016/j.procs.2023.10.349.

Personal Contribution

For Paper I, the author of this thesis contributed to the planning, execution, and writing of the research. For Paper IV, the author of this thesis contributed to the planning and execution of the research and was the main contributor responsible for writing the manuscript. For other included papers, the author of this thesis was the main contributor with regard to the planning, execution, and writing of the research.

ACKNOWLEDGEMENT

First and foremost, I would like to express my deepest gratitude to my supervisors, Johan Holmgren, Radu-Casian Mihailescu, and Jesper Petersson. Your continuous support, insightful guidance, availability, and invaluable expertise have been instrumental in shaping both my research and this doctoral thesis. Thank you for your commitment and mentorship throughout this journey.

I am also deeply thankful to my PhD examiner, Paul Davidsson, and the follow-up group members, Jan Persson, Nancy Russo, and Dario Salvi, for their constructive feedback and academic guidance, which have greatly contributed to improving my research and study plans. My gratitude also extends to Susanne Lundborg for her invaluable support during my PhD studies. I would like to thank my colleagues at the Department of Computer Science and Media Technology—especially my fellow PhD students—whose camaraderie, friendship, and shared experiences have been truly inspiring.

Finally, I would like to express my gratitude to my family and friends for their continuous encouragement and motivation. My deepest thanks go to my parents, who have always supported me unconditionally and laid the foundation for my success. Ali and Ahmad, I am grateful to have you as my brothers—thank you for all the fun times we shared in our childhood. My heartfelt appreciation goes to my wife, Faezeh. Your unconditional kindness, love, and unwavering support have been a powerful motivator throughout my PhD journey. And to my two-year-old daughter, Hanna, you are always a source of joy and inspiration—your smiles and laughter have brightened my days and brought balance to this demanding path.

LIST OF ACRONYMS

ANN	Artificial Neural Networks
CT	Computed Tomography
DES	Discrete Event Simulation
EMS	Emergency Medical Services
EV	Emergency Vehicle
MILP	Mixed-Integer Linear Programming
ML	Machine Learning
MRI	Magnetic Resonance Imaging
MSU	Mobile Stroke Unit
ORS	Open Route Service
OSM	Open Street Map
OSRM	Open Source Routing Machine
OSRM-CH	Open Source Routing Machine with Contraction Hierarchies
QGIS	Quantum Geographic Information System
RA	Regular Ambulance
SHR	Southern Healthcare Region
TIA	Transient Ischemic Attack
tPA	tissue Plasminogen Activator
WATT	Weighted Average Time to Treatment

PART I: COMPREHENSIVE SUMMARY

1. INTRODUCTION

Stroke is one of the leading causes of death and disability worldwide. Immediate treatment is crucial for improving a stroke patient's chances of recovery and reducing the social and financial burden of long-term care. Every minute that elapses from the onset of a stroke until the patient receives treatment affects the successful recovery of the patient. To minimize delays in delivering timely treatment to stroke patients, implementing well-established stroke transport policies is essential. One such policy involves using mobile stroke units (MSUs) alongside regular ambulances (RAs) in prehospital stroke care. An MSU is a specialized ambulance equipped with a computed tomography (CT) scanner and staffed by trained medical personnel who can diagnose and begin treating stroke patients directly at the scene, potentially reducing the time to treatment and improving patient outcomes.

The prehospital stroke care chain is inherently more complex than typical emergency medical services (EMS) care chains for other medical conditions. A key challenge in the prehospital stroke care chain is the variety of decisions that need to be made due to the availability of different emergency vehicles (EVs), that is, MSUs and RAs, which may be dispatched to provide service to the patient. Additionally, these decisions often need to be made with incomplete information; for example, determining whether to transport a stroke patient to the nearest hospital or a special clinic can be challenging without first performing a CT scan on the patient's brain.

Investing in an MSU service is a significant decision for healthcare providers, as it carries substantially higher costs than RAs. Therefore, such an investment is only justified if MSUs can deliver significantly better outcomes for stroke patients. The benefits of using MSUs and their impact on reducing time to diagnosis and treatment depend significantly on where and how they are placed. Optimization approaches can help determine the most suitable MSU placements based on different perspectives, that is, maximizing population coverage or ensuring equitable access to care. However, optimization alone may not capture all relevant aspects, such as variability in patient conditions, real-time traffic conditions, the dynamic availability of emergency resources, or regional healthcare differences.

Furthermore, to avoid unintended consequences, stroke transport policies must be carefully evaluated before implementation in real-world settings. As assessing such policies with real patients poses risks and practical difficulties, computer simulation emerges as a valuable tool. Simulation models offer a cost-effective and time-efficient means of assessing numerous policies under various configurations without compromising patient safety.

The purpose of this thesis is to contribute to earlier access to high quality stroke care and mitigate the impacts of stroke on society. The main aim is to develop optimization and simulation models to improve the analysis and planning of prehospital stroke care. To achieve this aim, we explored four objectives:

1. Develop an optimization model to identify optimal MSU placements that can provide efficient and equal access to stroke treatment for most inhabitants across a region.
2. Evaluate, through simulation, various stroke transport and dispatching policies, including the use of MSUs, for stroke patients requiring different treatments.
3. Introduce a modeling framework for constructing simulation models that reflect real-world EMS operations and support the analysis of different EMS policies.
4. Integrate dynamic travel time calculation and ML-based travel time estimation functionalities into the EMS simulation modeling framework to assess their impact on RA travel time estimations.

The findings from this thesis are intended to assist public health practitioners in refining existing stroke transport policies.

The remainder of this chapter is organized as follows: first, we present background information on stroke and stroke transport policymaking, followed by an overview of optimization and simulation modeling in healthcare transport. Next, we outline the problem statement and research questions used throughout the research. Finally, we conclude with an outline of the thesis structure.

1.1. Background

In this section, the fundamental concepts that form the basis of this thesis are described.

1.1.1. Stroke

Stroke is the second leading cause of death globally [1]. According to the Global Stroke Fact Sheet 2022 provided by the World Stroke Organization [1], one in four people over the age of 25 will experience a stroke in their lifetime. Every year, 12 million people suffer a stroke, resulting in 6.6 million deaths [1]. In Sweden, the Swedish Stroke Register (Riksstroke) [2] recorded approximately 20,000 cases of stroke in 2023, including both first-time and recurrent strokes. In the same year, there were around 4,000 cases in Sweden's Southern Healthcare Region (SHR), with 373 cases in Blekinge County, 760 in Southern Halland County, 409 in Kronoberg County, and 2,444 in Skåne County.

Stroke is also a major cause of long-term physical and cognitive disability, often leaving individuals unable to carry out daily activities [3]. This not only increases healthcare costs for patients but also places an emotional burden on their relatives while limiting their ability to work. Furthermore, stroke imposes substantial costs on society and public health due to the need for long-term care and rehabilitation.

A stroke is a disease that occurs when a blood clot or a hemorrhage in the brain either disrupts or reduces blood flow to the brain, leading to a lack of oxygen. This can cause sudden symptoms in the individual, such as difficulty speaking, paralysis or numbness (especially on one side of the body), loss of vision, trouble walking, dizziness, and severe headaches. Without continuous blood flow, brain cells die rapidly—it is estimated that two million brain cells are lost every minute during a stroke [4]. As a result, access to immediate treatment is critical to enhance the chances of successful recovery for stroke patients.

There are three main types of stroke: ischemic, hemorrhagic, and transient ischemic attack (TIA). Although the symptoms of these stroke types are often similar, each requires specific treatment. An *ischemic* stroke occurs when one or more blood clots reduce blood flow inside the brain, requiring treatments like thrombolysis or, in cases of larger clots, thrombectomy. Thrombolysis involves the intravenous administration of tissue plasminogen activator (tPA) to dissolve the clot and restore blood flow, while thrombectomy involves mechanically removing the clot using a specialized device at a special clinic [4]. A *hemorrhagic* stroke occurs when a blood vessel in the brain ruptures, causing bleeding into the surrounding tissue. Immediate treatment typically involves lowering blood pressure to prevent further bleeding. A *TIA*, or mini-stroke, occurs when blood flow is temporarily blocked by a blood clot, but the brain recovers fully after a short period. Treatment for TIA typically involves taking medications that reduce the risk of future strokes.

The type of stroke is usually identified through a CT scan or magnetic resonance imaging (MRI) of the brain, procedures that are generally only available in hospitals. Misdiagnosis, particularly failing to differentiate between ischemic and hemorrhagic strokes, can severely endanger the patient's health. For example, administering thrombolysis to a patient suffering from a hemorrhagic stroke could result in life-threatening complications.

Ischemic stroke (including TIA) and hemorrhagic stroke account for approximately 87% and 13% of all stroke cases, respectively, both globally and in Sweden [2]. Due to the higher prevalence of ischemic stroke, most stroke research focuses on developing effective treatment strategies for this type of stroke. In Sweden, among the ischemic stroke cases registered in 2023, a total of 81% did not receive treatment, either because patients did not arrive at the hospital in time for intervention or, after examination, it was determined that treatment was not needed at that moment. Of the remaining 19% who underwent recanalization (reperfusion) therapy, 11% received thrombolysis alone, 3% received a combination of thrombolysis and thrombectomy, and 5% underwent thrombectomy alone [2].

The most effective way to minimize the consequences of a stroke, including disability and death, is to deliver immediate treatment following an accurate diagnosis. In this regard, the term *golden hour* is proposed for ischemic strokes to emphasize the importance of initiating treatment within one hour of symptom onset, as this greatly increases the chances of satisfactory recovery [5]. Providing immediate treatment is often delayed mainly due to logistical challenges and the inability to diagnose the type of stroke quickly. These challenges may arise from a shortage of EVs in some areas or long travel times to the nearest hospital [6]. Additionally, as CT scanners are generally only available in hospitals, treatment cannot begin until after the patient has been transported to and diagnosed at an acute hospital.

It is important to note that the benefits of thrombectomy are typically greater than those of thrombolysis, especially for patients with large blood clots [5]. However, only a limited number of regional hospitals, that is, special clinics, offer thrombectomy, primarily due to the specialized resources and expertise required. As a result, patients living farther from a special clinic often experience longer delays in receiving thrombectomy compared to those living closer. For example, the SHR, which is the focus area of this thesis, is a vast region encompassing four counties—Blekinge, Southern Halland, Kronoberg, and Skåne—and 49 municipalities, with a total area of approximately 29,345 square kilometers and a total population of approximately 1.9 million people in 2023 [7]. As illustrated in Figure 1.1, the SHR has 13 acute hospitals equipped with CT scanners and 39 active ambulance sites. However, Skåne University Hospital Lund (Lund SUS), represented by a yellow diamond in Figure 1.1, is the only special clinic in the SHR capable of performing thrombectomy. In Figure 1.1, population density is shown by tiny colored squares, where darker red represents higher densities.

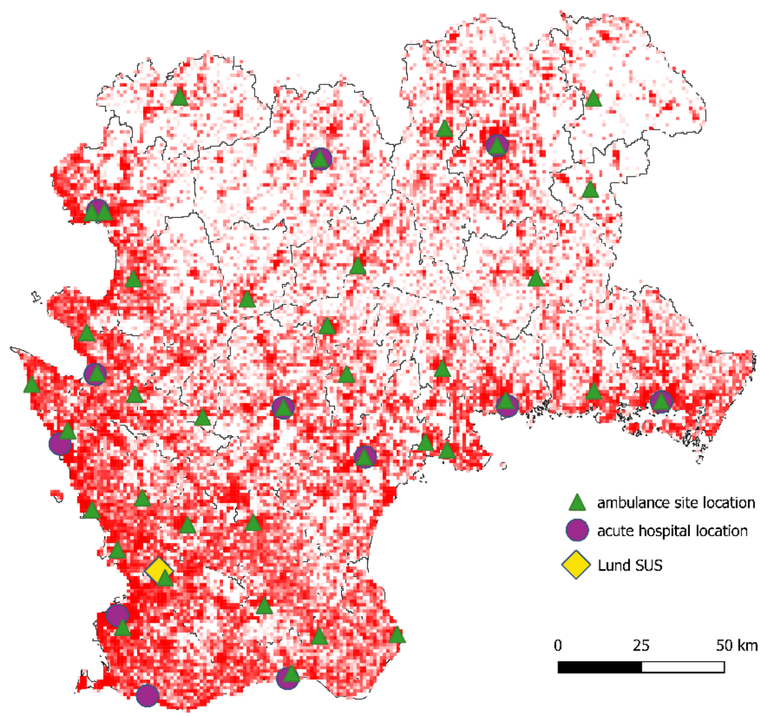


Figure 1.1. An overview of the SHR. Green triangles represent ambulance site locations, and purple circles indicate acute hospitals. The only special clinic for thrombectomy treatment in the SHR, Skåne University Hospital Lund (Lund SUS), is marked by a yellow diamond. Each square's color reflects population density, with lighter squares indicating lower density.

1.1.2. Stroke Transport Policymaking

A policy is a principle of action, typically framed as a set of guidelines, which can be used to make important decisions, including identifying different alternatives and choosing the best option based on their potential impacts within a specific domain. Policy assessment involves evaluating and predicting the potential impacts of planned policies on the target group, along with their social and economic outcomes, to support decision-makers [8]. In healthcare, policy assessment is essential for improving patient outcomes and enhancing the quality of services.

In this thesis, we evaluate *EMS* and *stroke transport policies* using the proposed models. In Figure 1.2, we illustrate how *transport policies* and *healthcare policies* are related to each other. Transport policies are the set of policies that concern the transport of goods and people. Healthcare policies refer to policies applied within the healthcare domain, including EMS policies and stroke policies. EMS policies focus on delivering emergency medical services for prehospital stabilization, treatment, and transport of severely ill or injured patients. Stroke policies, a subset of healthcare policies, are specifically applied within stroke care. A stroke transport policy is a type of stroke policy applied to the transportation of stroke patients, with the primary aim of improving prehospital stroke care by reducing the time to diagnosis and treatment. Stroke transport policies and EMS policies lie at the intersection of transport and healthcare policies, as represented by the orange area in Figure 1.2.

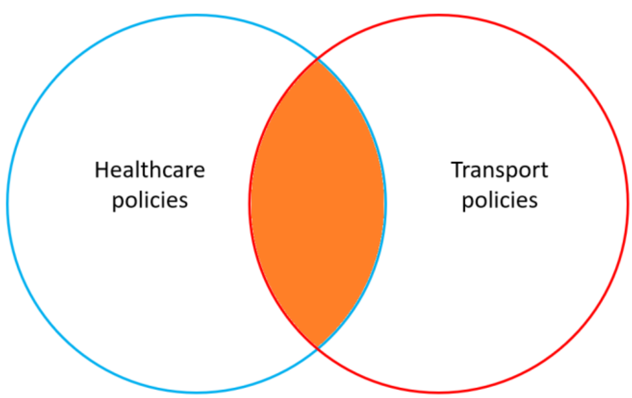


Figure 1.2. Illustration of the relationship between transport and healthcare policies, highlighted in orange, where the EMS and stroke transport policies lie at the intersection of these two policies.

To improve the care provided to stroke patients, all activities within the stroke care chain should be performed effectively. This includes, for example, activities related to travel time to the hospital, diagnosing the type of stroke, and monitoring the availability of resources. An optimal stroke transport policy for stroke patients is expected to lead to better stroke treatment and ensure that many patients receive timely and appropriate care. However, achieving the same level of service for all patients may not be feasible due to factors such as geographic limitations.

Stroke transport policies can include dispatch strategies, such as single and co-dispatch, which determine how EVs are deployed to stroke patients. Single dispatch involves sending one EV, while co-dispatch sends multiple EVs, such as an RA and an MSU, to the scene to expedite care.

1.1.2.1. Mobile Stroke Unit (MSU)

As with other emergency medical conditions, an RA is typically used in prehospital stroke care to transport patients to medical facilities. An alternative is the deployment of MSUs in prehospital stroke care [6], which offers the advantage of diagnosing and treating patients on site. The MSU concept was first proposed in 2003 and developed further in 2008 in Germany by Dr. Klaus Fassbender and his team, with the goal of optimizing prehospital stroke care [9, 10]. MSUs are specialized ambulances equipped with a CT scanner, a hematology analyzer, and other medical emergency tools. They are staffed by a team that typically includes a paramedic or nurse, a CT technician (such as a radiographer or radiologist), and a stroke specialist or neurologist, either onboard or via telemedicine.

Figure 1.3 presents an MSU and its interior, operated in Berlin, Germany. The CT scanner in the MSU enables the team to diagnose stroke patients by differentiating between ischemic and hemorrhagic strokes, allowing for the administration of intravenous thrombolysis at the patient’s location without the need for transport to a hospital. In addition to diagnosing ischemic and hemorrhagic strokes, the CT scanner in the MSU also facilitates the diagnosis of other conditions, such as brain masses and focal seizures.

The use of MSUs helps reduce the time to diagnosis and treatment by at least the amount of time required to transport and diagnose the patient at the hospital. Most relevant papers review the advantages of using MSUs for ischemic stroke patients, particularly focusing on those who require thrombolysis, which can be administered directly inside the MSU. For thrombectomy cases, the primary time-saving benefit lies in making an earlier decision to transport the patient to the appropriate destination; that is, instead of taking the patient to the nearest hospital, they can be transported directly to a special clinic for thrombectomy. However, the potential benefits of initiating blood pressure-lowering therapy for hemorrhagic stroke patients in the MSU have not yet been explored.

Studies on MSU operations in cities such as Berlin [6], Cleveland [11], Melbourne [12, 13], and Oslo [14] demonstrate that compared to using only RAs, the use of MSUs results in a substantial reduction in time to treatment for ischemic stroke patients [6, 11, 13] and provides long-term financial savings for society [12, 14]. Furthermore, the additional costs of operating an MSU are expected to be compensated by long-term savings in the care of stroke patients [10]. Norway is the only Nordic country where at least one MSU has been operational. Currently, Sweden does not have any operational MSUs.

Figure 1.3. An MSU in Berlin, Germany, and its side, including the CT scanner [15]. Used with permission, copyright MEYTEC GmbH.

MSUs are generally larger and heavier than RAs due to their advanced diagnostic equipment, which can affect their speed and maneuverability, especially in urban environments. However, their ability to provide immediate treatment, such as thrombolysis, compensates for these logistical challenges and can potentially improve the outcomes for stroke patients. These benefits highlight the importance of careful consideration when deploying MSUs in a region to ensure their effectiveness.

In this thesis and the included papers, the term “treatment” refers specifically to stroke-related interventions, including thrombolysis and thrombectomy for ischemic stroke patients and blood pressure-lowering therapy for hemorrhagic stroke cases. These treatments are central to the models we developed and are the focus of our analyses. As defined in our models, time to treatment refers to the time elapsed before initiating treatment for the occluded brain artery in ischemic stroke cases (either thrombolysis or thrombectomy). In the case of hemorrhagic stroke, treatment is focused on lowering blood pressure, which can only begin once the patient’s brain has been scanned. In our models, a stroke-specific treatment can begin once the diagnosis (CT scanning) is completed. We clarify that time to treatment is not allocated a specific duration in the models but rather represents the activity of administering the stroke-specific treatment.

MSUs were initially introduced to provide thrombolysis for ischemic stroke cases, as they make up the majority of stroke cases. While MSUs can administer thrombolysis for ischemic strokes on site, blood pressure-lowering therapy for hemorrhagic cases is not yet typically performed inside MSUs in real-life operations, as this capability is still under development. At acute hospitals, thrombolysis for ischemic stroke patients and treatments for hemorrhagic strokes, such as blood pressure-lowering therapy, can be provided. Additionally, all types of stroke treatments, including thrombectomy, are available at special clinics.

It is important to note that RA personnel also provide a wide range of medical care upon arriving at the scene, including basic first aid and general medical support. However, in the context of our models, we are specifically focused on the stroke-specific treatments mentioned above—thrombolysis, thrombectomy, and blood pressure-lowering therapy. Other forms of treatment, such as managing external injuries or general medical support, are not considered in our models, as they fall outside the scope of this study.

1.1.2.2. Single Dispatch Policy

In a single dispatch policy, only one type of EV is assigned to respond to an incident. For example, Figure 1.4A illustrates a single dispatch policy in which an RA transports a stroke patient to the closest hospital. As shown in Figure 1.4B, sometimes the patient may need to be further transported to a special clinic, leading to additional transport.

As mentioned earlier, an alternative to the conventional single dispatch policy in prehospital stroke care is the use of MSUs [6]. In Figure 1.4C, we present a single dispatch policy using an MSU. Here, the patient is diagnosed inside the MSU, and based on the diagnosis, transported directly to the nearest hospital or special clinic. MSU dispatch can enable quicker diagnosis and treatment. For cases involving large clots, where mechanical thrombectomy is often required, MSUs can transport patients directly to a special clinic without unnecessary delay [16].

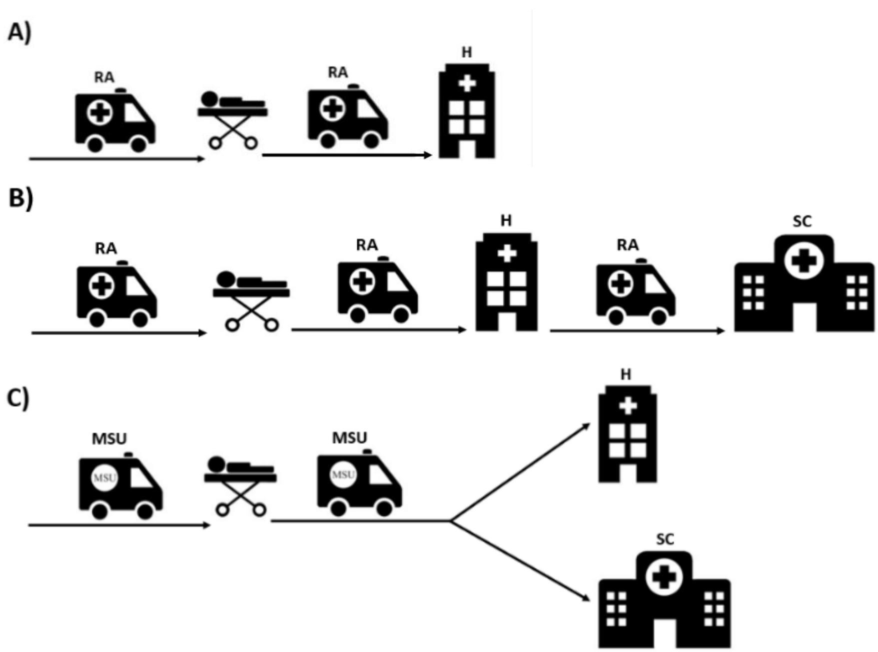


Figure 1.4. Single dispatch policies in prehospital stroke care for an acute stroke patient: A) using an RA to transport the patient to the closest hospital (denoted as H in the figure); B) using an RA to transport the patient to the closest hospital and then to a special clinic (denoted as SC in the figure) and; C) using an MSU to transport the patient to the closest hospital or a special clinic based on the diagnosis made inside the MSU. The graphs in this figure are inspired by Mathur et al. [17].

Below, we present the main flow of prehospital stroke care chain activities in the single dispatch policy for an ischemic stroke patient, from symptom onset to the administration of thrombolysis. We assume that both RAs and MSUs are available to provide service to the patient.

- A stroke occurs, and the emergency center receives an emergency call.
- An EV (RA or MSU) is assigned to the patient based on availability and predefined decision policies.
- *If an MSU is assigned for the operation:*
    - The MSU is dispatched from its current location to the patient's location.
    - The MSU arrives at the patient's location. The paramedics examine the patient, transfer the patient to the MSU, and diagnose the patient inside the MSU by performing a CT scan on the patient's brain.
    - The paramedics prepare and secure the patient for transportation and thrombolysis.
    - The thrombolysis is initiated inside the MSU.
    - Once thrombolysis is started, the paramedics decide to which hospital the patient should be transported based on the diagnosis and predefined policies.
    - The MSU departs for the destination hospital.
    - The MSU arrives at the hospital, and the patient is unloaded.
    - After unloading the patient, the MSU becomes available for the next operation. Alternatively, it may return to its station before becoming available for the next operation.
- *If an RA is assigned for the operation:*
    - The chosen RA is dispatched from its current location to the patient's location.
    - The RA arrives at the patient's location. The paramedics examine and transfer the patient to the RA.
    - Depending on the hospital selection policy, a destination hospital will be chosen for the patient.
    - The RA departs for the selected hospital.
    - The RA arrives at the hospital, and the patient is unloaded.
    - The patient is diagnosed at the hospital.
    - The thrombolysis will be initiated at the hospital.

- After unloading the patient, the RA becomes available for the next operation. Alternatively, it may return to its station before becoming available for the next operation.

As presented above, under a single dispatch policy involving an RA, the RA needs to travel to two distinct destinations (the patient's location and a hospital) before an ischemic stroke patient can receive thrombolysis: 1) driving from the RA's current location to the patient's location; and 2) driving from the patient's location to the selected hospital. However, when an MSU is assigned to an ischemic stroke patient, the patient can receive thrombolysis directly at their location. Additionally, the MSU can also assist hemorrhagic stroke patients, at least by reducing the time required for diagnosis at the hospital.

When both RAs and MSUs are available to be dispatched, a decision should be made regarding which type of EV should be assigned to provide service to a stroke patient. This decision can be made by analyzing the approximate time either to 1) get an EV to the patient's location, 2) diagnose the patient, or 3) initiate patient's treatment, as presented in detail in Paper IV of this thesis. Another typical decision in the prehospital stroke care chain is choosing a hospital to where the patient should be transported. This decision can be made by identifying either the closest hospital to the patient's location or the closest special clinic.

When an EV delivers a stroke patient to the acute hospital, the EV paramedics typically do not remain at the patient's bedside, and the EV leaves the hospital. Glober et al. [18] contribute a clinical-based study to examine the use of a new EMS policy, asserting that after delivering a patient with a suspected stroke to an acute hospital, the paramedics should remain at the patient's bedside until clinical assessment and diagnosis are completed, and they are given permission to leave the hospital. If the patient needs special treatment, the RA paramedics at the bedside transfer the patient using the same RA to the closest special clinic. The results demonstrate that the proposed policy reduces the time required to transfer patients in need of thrombectomy to a special clinic. However, the authors do not take into account the availability of the RA and its paramedics remaining at the hospital for other patients who may need the same RA simultaneously.

Prehospital stroke care chain activities in a single dispatch policy can typically consist of five transport types:

1. Travel from the EV's current location to the patient's location,
2. Travel from the patient's location to a hospital,

3. Travel from a hospital to an ambulance site,
4. Travel from the patient's location to an ambulance site,
5. Travel from one hospital to another, such as to a special clinic.

1.1.2.3. Co-dispatch Policy

In EMS, co-dispatching or a rendezvous approach refers to a situation where two or more EVs are simultaneously assigned for an emergency incident and collaborate until the patient receives appropriate treatment. For example, involving both an RA and an MSU in the transportation of a suspected stroke patient can provide more timely and effective care [17]. The involved EVs rendezvous at a predetermined meeting point to transfer the patient from one EV to another. Alternatively, aircraft such as helicopters can be involved when co-dispatching [17]. In addition, Walter et al. propose the idea of the Air-MSU concept as a tool to reduce treatment inequities for stroke patients in rural areas [19].

In co-dispatching, upon receiving a stroke-related call, one EV, such as an RA, is dispatched to the patient's location, while the other EV, such as an MSU, is sent directly to a predetermined meeting point. After picking up the stroke patient, the RA travels toward the meeting point, where the two EVs meet and transfer the patient. Then, the MSU transports the patient to an appropriate hospital based on the diagnosis made inside the MSU. Figure 1.5A presents this scenario of co-dispatching for a stroke patient, which is, for example, under operation in Northern Alberta, Canada [20]. In this scenario, the RA can instead transport the patient from the meeting point to the destination hospital so that the MSU becomes immediately available for an upcoming operation. Alternatively, depending on operational needs during co-dispatching, the MSU might travel directly to the patient's location while the RA is dispatched to the meeting point. The alternative scenario of co-dispatching is illustrated in Figure 1.5B.

In co-dispatching for stroke care, both scenarios offer advantages in reducing the time to diagnosis and treatment compared to a traditional RA dispatch policy. In the first scenario, the MSU travels directly to the meeting point without needing to go to the patient's location. This approach shortens the MSU's operation time, allowing it to be more efficient and available for other emergencies. In the second scenario, the MSU becomes immediately available for subsequent emergency operations after the patient is transferred to the RA at the meeting point.

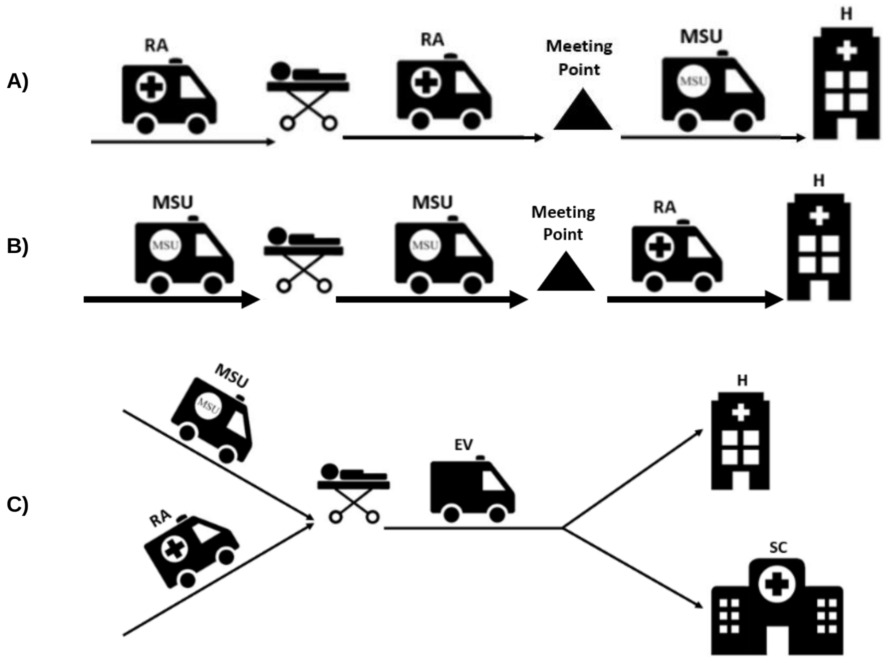


Figure 1.5. Co-dispatch policy in prehospital stroke care for an acute stroke patient. A) co-dispatching, where an RA is dispatched to the patient's location. Thereafter, the RA drives the patient to the meeting point. In the meantime, the MSU is dispatched directly to the meeting point. At the meeting point, the patient is transferred to the MSU. Based on the outcome of the patient's diagnosis inside the MSU, the patient will be transported to an appropriate hospital by the MSU; B) in contrast to A, the MSU drives toward the patient's location while the RA drives directly to the meeting point; and C) co-dispatching, where both an RA and an MSU are simultaneously dispatched to the patient's location, and then based on the outcome of the patient diagnosis inside the MSU, the patient will be transported to either an acute hospital or a special clinic using an EV.

In the co-dispatching of an RA and an MSU, a meeting point refers to a predetermined or dynamically calculated location where the two EVs converge to transfer or collaborate on the care of a stroke patient. Strategically determining a meeting point during co-dispatching is crucial for providing timely service [13]. Safe and practical locations within road networks must be considered as meeting points for patient transfers. For example, factors such as road type, connectivity, accessibility, one-way streets, practical transfer locations, and wireless signal strength [17, 21] should be evaluated when selecting an appropriate meeting point. A few studies in prehospital stroke care have explored the problem of identifying safe and practical locations for transferring a patient, with suggestions including choosing the meeting point as the patient's location [13], approximately half the distance from the MSU location [17], midway between the patient's location and the MSU location [20, 22], or parking lots and safe roads [21].

As an example of considering a patient's location as a meeting point, Figure 1.5C illustrates co-dispatching where both an RA and an MSU are simultaneously dispatched to the patient's location. Thereafter, depending on the diagnosis outcome, the patient is transported to either an acute hospital or a special clinic using one of the involved EVs.

1.1.2.4. Policy Assessment

Policy assessment in stroke care contributes to establishing a practical strategy to, for example, expedite the quick response to stroke incidents and reduce the time from when stroke symptoms have appeared until the patient receives treatment. It is generally complicated to assess policies manually, for example, due to the large number of different cases that might occur. Complex policies can instead be analyzed through computer-aided modeling, which enables the modeling and prediction of the behavior of real systems on different levels, for example, micro-level or macro-level, without putting the health of the people at further risk. The use of modeling, in the form of optimization and simulation, can be used to assist policymakers in building more helpful policies on the basis of the strengths and weaknesses of the model results. For example, assessing policies related to prehospital stroke care can lead to an optimal stroke transport policy for stroke patients.

As an example of studying stroke transport policies using only RAs, Al Fatah et al. [23] develop an agent-based simulation model to evaluate two stroke transport policies concerning where to transport potential stroke patients for diagnosis, that is, the *nearest hospital policy* and the *nearest hospital towards the stroke center policy*, which respectively analyzes whether patients should be transported to the closest hospital or to the closest hospital in the direction of a special clinic. The simulation results show that patients requiring special treatment, that is, thrombectomy, benefit from being transported to the closest hospital in the direction of a special clinic, while the patients who do not need special treatment benefit from being transported to the closest hospital. In addition, the authors conclude that since the time required to transport the patient to the hospital for diagnosis is higher for the *nearest hospital towards the stroke center* policy, the *nearest hospital* is most likely a better policy to be implemented in the SHR.

When introducing EVs in a geographic region, it is essential to consider the effects of how they are placed so that they can provide maximum benefit for all inhabitants regardless of where they live. Previous studies in EMS typically propose two perspectives for how to place EVs in a region: efficiency and equity [24, 25]. Efficiency refers to optimally placing EVs to cover as many inhabitants/patients as possible, ensuring they can receive treatment in the shortest time. Equity focuses on placing EVs in a region to promote equal care for all inhabitants/patients, particularly those who currently live far from medical centers or ambulance sites. EVs are typically placed in either urban or rural areas to address efficiency or equity, respectively. Accordingly, Dahllöf et al. [26] study the effects of MSU placement in Skåne County, Sweden, in light of the efficiency and equity perspectives.

1.1.3. Modeling

Modeling is the process of constructing a model of a real system, where a model is a simplified representation or abstraction of the key characteristics or behaviors of the system. A model can be used for purposes such as optimization or simulation, and it typically contains data and components that are linked in a specific way. The model should be simple yet provide a sufficiently accurate approximation of the system it represents. It can serve as a valuable tool to support the authorities in the decision-making processes.

The use of modeling in healthcare is increasing mainly due to the availability of data, the availability of faster computers, and growing awareness in the medical field of how modeling can support decision-making and improve services. EMS models are primarily constructed using optimization and simulation approaches. Optimization models aim to identify optimal solutions for specific problems without violating defined constraints. Simulation models are particularly useful for studying healthcare systems over time and analyzing different scenarios and policies.

1.1.4. Healthcare Transport Optimization

Mathematical optimization is the use of mathematical models and methods to identify the best possible solution for a decision problem under consideration. An optimization model consists of a number of decision variables that are used to minimize or maximize the value of an objective function, and a number of constraints that restrict the values of the decision variables. The objective of an optimization model is to find the values of the decision variables that yield the best value of the objective function while fulfilling the constraints defined on those variables [27]. The primary objective in healthcare transport optimization is to ensure patients receive timely care. The purpose of an optimization model in EMS can be, for example, to identify the optimal placement of EVs in the study region.

One commonly used approach to represent a mathematical representation of an optimization problem is mixed-integer linear programming (MILP), which formulates problems as sets of linear equations and inequalities involving both continuous and integer variables, effectively capturing the key characteristics of the problem it represents [27]. MILP is suitable for linear optimization problems involving integer variables, such as determining whether to place an EV at a specific location. Formulating problems using MILP enables the use of more efficient optimization methods that systematically search the solution space to find the optimal solution.

One approach for solving optimization problems is exhaustive search, a brute-force method that evaluates all possible solutions to find the optimal configuration of resources. While the exhaustive search guarantees to identify the optimal solution, it does not employ any advanced search techniques and can become computationally infeasible for large problem instances due to the exponential growth in the number of possibilities. Alternatively, (meta)heuristics are approximate algorithms that aim to find feasible, but not necessarily optimal, solutions within a reasonable amount of time. Heuristics trade off some level of accuracy for faster computational times, making them suitable for large-scale problems where finding the exact optimal solution is not possible [27]. Common heuristics applied in EMS optimization include greedy algorithms, evolutionary algorithms, and simulated annealing, each offering different ways to explore the solution space. It should be noted that when solving an optimization problem using an exhaustive search or heuristic approach, we may not require the same level of mathematical formulation as in an MILP model for the same problem.

MILP models are typically solved using exact optimization solvers, such as Gurobi and CPLEX, which guarantee optimality under certain conditions, ensuring that the best possible solution is found. Solving MILP models using these solvers allows for systematic exploration of the solution space, which can potentially reduce computational time compared to solving the same problem directly using an exhaustive search algorithm. However, solving MILP models using optimization solvers can become computationally expensive for large problem instances, as they may require a significant amount of time and high-performance computing resources. Despite this, recent advances in solver technologies and decomposition methods have made MILP increasingly practical for applications like healthcare transport optimization. Previous studies in prehospital care and EMS have applied mathematical optimization and MILP to address challenges such as RA routing and placement, fleet allocation, crew scheduling, and resource distribution, ultimately improving patient outcomes and optimizing resource utilization [24, 28-30].

MILP models can be also solved using heuristics when exact solutions are infeasible due to computational constraints. Solving MILP models using optimization solvers is preferable to heuristics when precision is critical as these solvers provide guarantees of optimality. In contrast, heuristics are often computationally faster and provide approximate solutions. Solving MILP models using solvers is also preferable to exhaustive search because they employ advanced search algorithms to systematically explore the solution space without evaluating every possible solution, thereby potentially reducing computational time.

Buying and operating MSUs in a region is a significant investment, making it essential to place them optimally to ensure they benefit all inhabitants within the region. A few studies in the literature focus on optimizing MSU placement, highlighting benefits for urban residents (efficiency perspective) [31, 32] or rural residents (equity perspective) [17]. Most prior studies prioritize efficiency, aiming to place an MSU to provide treatment to as many inhabitants/patients as possible in the shortest time. Equity, addressed in fewer studies, focuses on ensuring equal access to stroke treatment, particularly for inhabitants with the longest times to treatment. For example, Dahllöf et al. [26] propose an exhaustive search method, referred to as the expected value optimization approach, to identify the optimal placement for an MSU in Skåne County, Sweden, aiming to evaluate the potential benefits of MSU placement from both efficiency and equity perspectives. Abid et al. [33] propose an approach using a genetic algorithm to optimize the placement of MSUs in the SHR. Their approach focuses on minimizing the time to treatment by identifying the most suitable ambulance sites for placing MSUs.

Achieving high levels of efficiency and equity simultaneously in the MSU placement problem is inherently challenging due to the conflicting nature of these perpectives. This conflict arises because each perspective is biased toward a specific group of inhabitants: efficiency benefits densely populated urban areas and equity prioritizes rural regions with longer time to treatments. Some recent studies in EMS aim to address this bias by proposing trade-offs between efficiency and equity for the optimal placement of RAs in a region.

1.1.5. Healthcare Transport Simulation

Simulation can be described as the imitation of the operation of a real or fictive system, which is often carried out over time. Simulation modeling is used to study a model numerically to predict its performance in real-world situations and to retrieve useful output from it [34]. In general, the main components of simulation models are system entities, input parameters, performance measures, and functional relationships. During the simulation, a set of entities are modeled, and the actions and connections between these entities are studied, sometimes over a specific period of time. The use of simulation allows for the modeling of complicated relations between different entities, particularly as it may be difficult to model these relationships in a mathematical model. In addition, simulation makes it possible to evaluate the performance of a system under different setups over long periods of time and assess decision policies before being implemented in the real system. Therefore, simulation helps eliminate the need to perform costly and time-consuming experiments on real systems.

Simulation models can be categorized as either deterministic or stochastic, depending on the presence of randomness in the model. A *deterministic* model contains no random variables and will produce the same results every time when executed with a particular set of inputs. In contrast, a *stochastic* model includes inherent randomness, where one or more input variables are controlled by probability distributions. Most simulation models are stochastic, meaning that at least one input or output variable is influenced by randomness [34]. This randomness allows stochastic models to better capture the uncertainty and variability found in real-world systems, such as patient arrival times, treatment durations, or traffic conditions in a healthcare setting. To assess the effect of stochasticity in a simulation model, the model can be run multiple times under the same conditions, and the variation in the results can be analyzed. Due to the random nature of some variables, the outcomes are expected to vary slightly between runs.

Simulation models can be classified into microscopic, macroscopic, and mesoscopic, depending on the level of detail in how entities and their interactions are represented. A *microscopic* simulation model is able to model entities and their interactions on the individual level, capturing detailed behaviors. *Macroscopic* models make use of aggregate data representing the individual entities on a larger scale, for example, at the regional level, and do not focus on individual behaviors. Macroscopic models often do not simulate individual events over time. *Mesoscopic* models bridge the gap between microscopic and macroscopic models. Mesoscopic models provide a high level of detail for entities but model their behaviors and interactions at a lower level of detail. Macroscopic and mesoscopic models are typically less computationally expensive than microscopic models, making them more suitable for studying large-scale systems. Stochastic models are more commonly found in mesoscopic and microscopic models, as they allow for capturing individual variability and randomness in system behaviors.

A wide range of simulation models are used in healthcare to support decision-making, largely due to their ability to analyze complex healthcare systems. Simulation models are applied to various areas in healthcare, including staff scheduling, allocation of human resources, patient flow management in emergency rooms, hospital bed utilization, and the management of prehospital healthcare services [35-39]. In particular, Keshtkaran et al. [40] classify the potential simulation modeling approaches in the stroke care systems into simulation modeling for stroke care operations improvement, economic analysis, public policy, and clinical applications.

Discrete Event Simulation (DES) and agent-based simulation are the most popular simulation modeling approaches that are used in EMS, including stroke transport [23, 41]. DES is a simulation modeling paradigm used to model the behavior of a system as a sequence of discrete events that occur over time. It only models the points in time where events occur and where the system's state can change. In contrast, in the fixed-increment time approach, the simulation is advanced using fixed-length time steps, where each time step is explicitly simulated. DES is a form of micro-simulation model, where individual entities' behaviors and interactions are explicitly modeled over time, typically allowing more realistic modeling. For example, in an EMS model, the DES model enables the study of the behavior of individual patients and EVs. Given that the macro-level simulation paradigm models the average of the characteristics for the whole population, it is not possible to consider the effects of unexpected and simultaneous events using the macro-level paradigm.

An agent-based model is a modeling approach with a collection of agents used to simulate the activities and interactions of individual entities in order to analyze the behavior of a system. Each agent is a program entity with a high level of autonomy, which means it is capable of, for example, making independent decisions and taking actions in response to its surroundings [42]. Some agent-based simulations share similarities with DES as they can use discrete event logic to manage state changes over time.

A modeling framework is a tool with reusable components that typically provides a general structure to build models of a real system. In healthcare, modeling frameworks are used, for example, to build models for home healthcare service delivery decisions for the aging population [43] and for resource allocation [44]. Modeling frameworks generally provide a graphical user interface to facilitate the construction of models, for example, using an activity diagram. An activity diagram or workflow diagram is a behavioral diagram that describes the workflow of activities from a starting point to an ending point, typically with the help of decisions. As activities and policies are modeled on the general level in the framework, the framework can be used to build specific models only by specifying the input data and a care chain specification.

1.1.5.1. Synthetic Stroke Population

A micro-level simulation model for prehospital stroke care analysis is driven forward by a population of patients, where each patient can be characterized by specific attributes such as symptom onset (hour, day of the week, day of the year), patient's location (coordinates) in the considered geographic region, age, sex, symptoms, diagnosis, and preferred treatment. The population can be either real (representing actual historical patients) or synthetic (generated using available statistics). From a modeling perspective, whether the population is real or synthetic is less critical than ensuring that the data accurately reflects the intended use case. While real data inherently reflects actual scenarios more accurately, synthetic data is generated using statistical and demographic distributions to approximate real-world conditions as closely as possible.

The use of a synthetic population is typically preferred for privacy reasons, as real individual-level patient data is often inaccessible. In addition, when the real dataset is insufficient to run the simulation model, synthetic populations provide a practical alternative by creating datasets that resemble real-world distributions. A synthetic population generation model can produce various populations of synthetic patients, where each set includes patients with different attributes. This diversity facilitates the exploration of different scenarios and supports sensitivity analyses by varying patient attributes such as location, symptoms, and treatment preferences across simulation runs, thereby improving the model's reliability and the robustness of simulation outcomes. As stroke patients and their geographic locations vary annually, synthetic populations capture these variations, ensuring the model remains robust across different scenarios.

Building on this foundation, related studies in stroke care mainly use Monte Carlo simulation [23, 45] or Poisson distribution [46-48] to generate synthetic stroke populations. Monte Carlo simulation generates synthetic populations by probabilistically assigning attributes (for example, incident time, age, health status) to individuals based on real-world data distributions. A Poisson distribution can be used to generate synthetic populations by modeling the occurrence of events (for example, stroke incidents) over time or space, assuming that these events occur independently and at a constant average rate.

Al Fatah et al. [23] generate a synthetic stroke population using aggregate stroke patients' data, and they use Monte Carlo simulation to sample patient attributes such as the patient's age, address, incident time, and type of stroke. Wang et al. [48] and Aked et al. [47] investigate age and disparities in stroke incidences and case-fatality over time by creating a synthetic population, assumed to be Poisson distributed, attributed by age, ethnicity, and sex. In addition, Alassadi et al. [46] create a synthetic population using the socio-demographic census data, real stroke patients data, and travel data to study the travel patterns of suspected stroke patients. For each generated patient in the synthetic data, they sample the patient's age group, municipality, and stroke incident time using a non-homogeneous Poisson distribution.

1.1.5.2. RA Travel Time Estimations in EMS Simulations

Accurately calculating RA travel times is a key step in EMS simulations, as it directly impacts resource allocation and response strategies. An RA travel time estimation predicts the time required for an RA to travel from one location to another, such as from an emergency scene to a hospital, and it is crucial for ensuring the reliability of these simulations. Factors such as traffic conditions, time of day, day of the week, weather, and the characteristics of the road network can significantly impact travel times. Neglecting to account for these variables can result in inaccurate estimations; for example, the geographically nearest RA may not always be the fastest to reach a patient due to traffic or road obstructions at a given time. In time-sensitive cases like that of a stroke, precise travel time estimation is vital for selecting the optimal EV and hospital, potentially saving valuable minutes and improving patient outcomes.

When RAs are under operation, typically with lights and sirens, they often exceed the speed limit compared to non-emergency vehicles, such as passenger cars, and benefit from right-of-way privileges and reduced traffic delays [49, 50]. However, the travel times calculated by routing engines, such as Open Source

Routing Machine (OSRM), OpenRouteService (ORS), and Valhalla, typically represent average times for passenger cars. These routing engines, which are widely used in EMS and stroke simulation models, provide unlimited requests when deployed locally [51]. However, their outputs may differ from the actual travel times of RAs during emergencies.

RA travel times can be provided to simulation models using two approaches: pre-calculated or dynamic. The choice between these approaches depends on factors such as simulation complexity, data availability, and the need for real-time accuracy. Pre-calculated travel times are generated using routing engines and stored in advance using an origin-destination matrix for predefined routes in a geographic region. Pre-calculated travel times are static and cannot account for real-time traffic changes. Such a travel time data preparation approach is complex and time-consuming. However, this approach may be computationally efficient during simulations, as the travel times are already available, reducing the need for dynamic and real-time calculations.

Dynamic travel time calculation functionality generates travel times instantly during the simulation using routing engines such as ORS and OSRM. OSRM with Contraction Hierarchies (OSRM-CH) is particularly effective due to its speed and scalability, making it well-suited for dynamic travel time calculations in simulations [51]. This functionality enables the calculation of travel times between any two new locations during simulations, whereas the pre-calculated approach is restricted to predefined locations. Using this approach allows to better align the simulation outcomes with actual EMS operations. However, dynamic calculations may require more computational resources, thus potentially increasing the simulation run times. If we use online services such as ORS Web API for dynamic and real-time travel time calculations during simulations, they can account for real-time traffic conditions. However, these services typically impose a limit on the number of requests allowed within a specific period [51].

The travel times calculated using routing engines can be directly used as RA travel times in simulations. To account for the faster driving speeds of RAs during emergencies, a driving adjustment factor can be applied to adjust travel times calculated by routing engines. Alternatively, machine learning (ML) approaches can provide a more accurate way to estimate RA travel times by identifying patterns in historical and real-time data. ML encompasses a range of techniques that enable computers to analyze complex variations, such as traffic conditions, road characteristics, and RA speeds during emergencies. While ML-based methods offer more accurate estimation compared to using routing engines alone, they require large datasets for training and are more computationally expensive to implement. ML-based approaches are applied to estimate vehicle travel time in simulations [52] and to estimate RA travel times [53-55]. Previous research indicates that three ML methods—polynomial regression, XGBoost, and artificial neural networks (ANN)—perform particularly well in estimating travel times [53].

1.2. Problem Description

As mentioned earlier, recent studies both at the clinical- and computer-based levels indicate that the use of MSUs has the potential to reduce the time to treatment for stroke patients. The aim of this thesis is to study how simulation and optimization can be utilized to obtain improved analysis and planning of prehospital stroke care. As discussed earlier, optimization and simulation are valuable approaches that enable the study of prehospital stroke care and transportation. Optimization can be used to identify the optimal location for MSUs in a region depending on the intended perspectives and constraints. Simulation can be used to assess different policies related to prehospital stroke care and transportation before being applied in real-world situations.

Recent research has studied the optimal placement of an MSU in a geographic region, where the main focus is on investigating the impact of optimal placement of an MSU for inhabitants of urban (efficiency perspective) or rural areas (equity perspective). However, none of these prior research studies concentrate on the optimal placement of MSUs in a larger region, considering both of the mentioned perspectives. In addition, some studies in EMS use optimization modeling to make a trade-off between efficiency and equity for the RA placement problem. We have identified the need to propose an optimization model that makes a trade-off between efficiency and equity for optimal MSU placements in a region so that all inhabitants have equal and efficient access to stroke treatment regardless of where they live.

Previous studies in EMS have applied mathematical optimization to solve problems related to EV routing and placement, fleet allocation, crew scheduling, and resource distribution, leading to improved patient outcomes and more efficient resource use [24]. Although mathematical optimization has proven effective for RA placement, no existing work explicitly addresses the mathematical formulation of the MSU placement problem.

Simulation provides a valuable tool for assessing stroke transport policies on different levels, for example, macro-level or micro-level. Prior studies mainly focus on using simulation modeling to investigate the advantages of deploying an

MSU in densely populated areas or using microscopic simulation models to assess stroke transport policies with only RAs. Given these focused areas in prior research, our goal is to broaden the scope by employing simulation to analyze the potential benefits of using MSUs across larger regions, including both urban and rural areas, such as the SHR, and to analyze stroke transport policies involving MSUs for patients requiring different types of treatments. Macroscopic models are particularly useful for providing a high-level overview of policy impacts by estimating emergency metrics such as the average time to treatment for different parts of a region. This makes them well-suited for assessing regional strategies where population-wide effects are the primary concern.

In contrast, microscopic simulation models, in the form of DES, help us to evaluate stroke transport and dispatching policies in more detail, as they enable the study of individual entities, such as patients and EVs, over time. A key advantage of a DES model is its ability to model the sequence of activities and interactions between entities in dynamic and time-dependent systems, providing a more realistic representation of EMS operations. This allows for incorporating stochasticity, such as simultaneous incidents, and evaluating policies at an operational level with greater precision. In addition, a micro-level simulation model needs to be fed by a population of stroke patients. If real patient data is inaccessible or insufficient, we can instead use a synthetic stroke population, enabling diverse scenario testing and sensitivity analyses.

Constructing simulation models is often a complex and time-consuming task. A modeling framework can simplify the creation of DES models in EMS, thereby reducing the effort required for model construction. However, no existing framework in healthcare provides a general structure to fully capture the complexities of the EMS care chain, especially in the context of prehospital stroke care. By incorporating various decision policies, such as EV selection, dispatch type selection, and hospital selection, an EMS simulation framework can enable detailed analysis and the optimization of EMS operations.

Preparing and using pre-calculated travel data in EMS simulations can be time-consuming and limit the flexibility of transport simulation models. In contrast, integrating a routing engine into an EMS modeling framework enhances its capability to calculate travel times dynamically between any two geographic locations, eliminating the need for pre-calculated data and aligning simulation outcomes more closely with real-world EMS operations. Given the importance of dynamic travel times in EMS simulations, we identified the need to integrate a locally deployed routing engine into the existing simulation modeling framework to evaluate the effectiveness of dynamic travel time calculations during EMS simulations.

Accurate travel time estimation is crucial for the reliability of EMS simulations, as it directly impacts decision-making, resource allocation, and patient outcomes. While conventional routing engines are typically used to estimate RA travel times, they are primarily designed for passenger cars and do not accurately capture the travel patterns of RAs. In addition, pre-calculated travel times, while accurate, do not account for variations in traffic conditions at different times of the day. In contrast, dynamic travel times provide a real-time snapshot of current traffic patterns but become impractical for long-term simulations covering days, months, or years. An ML-based approach can bridge this gap by learning from historical data and capturing traffic patterns over time, making it a suitable solution for correcting the RA travel time estimations. Nevertheless, the benefits of using an ML-based approach in EMS simulations remain largely unexplored. For this reason, we identified the need to include an ML-based RA travel time estimation functionality in the established simulation modeling framework.

1.3. Research Questions

The main aim of this thesis is to improve the analysis and planning of prehospital stroke care by means of optimization and simulation. The focus is on assessing stroke transport policies related to involving MSUs in prehospital stroke care. To fulfill the aim of the thesis, the main research question is formulated as follows:

RQ. How can optimization and simulation modeling be used for improved analysis and planning of prehospital stroke care in the context of MSUs?

The main research question is addressed through four sub-research questions, as outlined below. Each sub-research question corresponds to a distinct objective, and the overall aim of this thesis will be achieved by fulfilling these objectives.

RQ1. How can optimization be used to determine the optimal placement of MSUs to provide equal and efficient access to stroke treatment?

In prehospital stroke care, optimization can help identify the best MSU locations considering constraints and different perspectives, such as efficiency and equity. RQ1 focuses on developing an optimization model that incorporates the trade-off between efficiency and equity for the optimal placement of MSUs in a region. It is important to examine the benefits of MSU placement for all inhabitants of the study region concerning each perspective and the trade-off between them. As efficiency and equity each have inherent biases toward specific groups of inhabitants, considering this trade-off is crucial to ensuring equal and efficient access to stroke treatment for most residents of a region. This question also focuses on formulating a mathematical model that captures the essential characteristics of the MSU placement problem. This model will integrate multiple constraints and an objective function to ensure that MSUs are positioned optimally given the operational needs and allocation perspective.

To address RQ1, in Paper II, we propose a trade-off function balancing the efficiency and equity perspectives for the optimal placement of MSUs in a geographic region, which is solved using an exhaustive search algorithm. Building on this, in Paper V, we present a mathematical model formulated using MILP to optimally place MSUs in a geographic region, where the objective function also incorporates the trade-off perspective.

RQ 2. How can simulation be used to evaluate stroke transport policies?

This research question explores the use of simulation to study and assess different stroke transport policies, in particular, the policies related to the use of MSUs. Research shows that a macro-simulation model can be helpful to achieve the average time to treatment for the population of a geographic region. In addition, a micro-level simulation model, for example, based on the DES paradigm, can be used to assess stroke transport policies in a geographic region and study the behaviors of EVs and patients. RQ2 is addressed in Paper I, where a simulation model is introduced through an average time to treatment estimation model that can be used to analyze the potential benefits of placing MSUs on the macro level. In addition, RQ2 is answered in Paper III, where we propose a micro-level DES model to study the advantages of using MSUs in the SHR by assessing different stroke transport policies.

RQ 3. What is an appropriate modeling framework for constructing simulation models to analyze real-world EMS operations?

The objective of RQ3 is to create a modeling framework for constructing DES models that analyze different EMS policies and represent real-world EMS operations. To address RQ3, we introduce a modeling framework in Paper IV that serves as a tool with reusable components, simplifying the process of building DES models for EMS policy analysis. This framework can be applied to construct EMS simulation models for various medical conditions, including stroke. Additionally, in Paper VI, we extend the framework by incorporating various dispatch policies, such as co-dispatch, to enhance its generalization in reflecting real-world EMS operations.

RQ 4. How does integrating dynamic travel time calculation and ML-based travel time estimation into the EMS simulation modeling framework impact RA travel time estimations?

Travel time estimation is an integral component of EMS simulations, as accurate travel times are essential for reliable and realistic modeling. Integrating dynamic travel time calculations directly into a simulation framework can make simulations more adaptable and reflective of real-world scenarios. Additionally, ML approaches can help determine actual driving times along a route or correct calculated travel times by routing engines. RQ4 explores how the use of dynamic and ML-based travel time estimations can influence the accuracy and adaptability of EMS simulations. To address this RQ, in Paper VII, we integrate a locally deployed routing engine into the simulation framework to enable dynamic travel time calculations and assess the benefits of this approach in emergency response scenarios. In Paper VIII, we also incorporate an ML-based travel time estimation module into the framework to evaluate how it impacts RA travel time accuracy and overall model reliability in EMS simulations.

1.4. Thesis Outline

The current thesis is divided into two parts: the first part provides a comprehensive introduction, and the second part consists of eight publications that form the basis of this thesis.

The remainder of the first part is organized as follows: Chapter 2 describes the research method used to address the research questions. Chapter 3 details the core contributions of the research and their connections to the research questions. Finally, Chapter 4 summarizes the main conclusions and suggests directions for future work.

2. RESEARCH METHOD

This chapter describes the research method applied throughout this thesis, where design science was adopted to guide the creation, evaluation, and application of artifacts that address the identified research questions. Design science aims to create innovative artifacts to expand human knowledge and to evaluate these artifacts in solving and addressing identified problems and research questions. An artifact is an object created by humans to solve real-world problems that people encounter in practice [56]. Using design science, we created five artifacts to fulfill the objectives of the thesis, addressing the identified research questions (RQs):

i. an optimization model (Paper II) without mathematically formulated constraints for tuning the trade-off between efficiency and equity in the MSU placement problem (RQ1);
ii. a mathematical optimization model in the form of an MILP (Paper V) to represent the MSU placement problem, which has the same objective as the first artifact (RQ1);
iii. a macro-simulation model (Paper I) for estimating the average time to treatment across a geographic region, enabling the exploration of the potential benefits of using MSUs (RQ2);
iv. a micro-simulation model (Paper III) for assessing stroke transport policies and studying the behaviors of patients and EVs (RQ2); and
v. a simulation modeling framework (Paper IV) to simplify the construction of simulation models for EMS policy analysis (RQ3). The framework was further extended in Papers VI, VII, and VIII to support different dispatch policies, dynamic travel time calculations, and ML-based RA travel time estimations (RQ3 and RQ4).

According to March and Smith [57], information technology artifacts created through design science can be classified into four categories: constructs, models, methods, and instantiations. *Constructs* are symbols and concepts used to formulate problems and potential solutions. *Models* represent practical problems and possible solutions built from constructs. *Methods* are the processes that are performed to solve problems. *Instantiations* demonstrate an artifact within its context, utilizing constructs, models, and methods to show the feasibility and effectiveness of the design process and the artifacts created [57]. The five artifacts developed in this thesis fall under the model category, as they represent formalized structures for addressing practical problems in prehospital stroke care.

It should be noted that the extension of the framework in Papers VI and VII falls under the model category, and it incorporates new dispatch policies, scenarios, and dynamic travel time calculation functionality into the framework, enhancing its ability to represent practical problems and potential solutions. In addition, the integration of ML-based approaches for estimating travel times into the framework (Paper VIII) introduces new procedures for estimating RA travel times, placing it within the method category.

As proposed by Peffers et al. [58], the research process in design science is typically divided into six main activities: problem identification and motivation, objectives definition, artifact design and development, artifact demonstration, artifact evaluation, and communication. In the following, we briefly outline how these six activities were applied in this thesis.

Activity 1: Problem identification and motivation. In the first activity, we investigated and analyzed various aspects of the problem through related work, group discussions, and consultations with domain experts. The insights and knowledge gained from addressing the research questions led to the problem formulation. In addition, we identified existing gaps in the research area that this research aims to address (see Section 1.2).

Activity 2: Objectives definition. In this second activity, we identified and defined the requirements and key components needed for designing and developing the research and outlined the possible solutions for the problem in the form of artifacts. We also reviewed optimization and simulation models developed in the EMS and prehospital stroke care domains.

Activity 3: Artifact design and development. During the artifact design and development process, we reviewed relevant studies to ensure that our artifacts aligned with established frameworks and care chain activities in EMS. To accomplish this activity, we created an optimization model without mathematically formulated constraints (Paper II), an MILP model (Paper V), a macroscopic simulation model (Paper I), a DES model (Paper III), and a modeling framework (Paper IV), which was extended to support different dispatch policies as well as dynamic and ML-based travel time estimation functionalities (Papers VI, VII, and VIII).

Activity 4: Artifact demonstration. The artifact demonstration activity precedes the evaluation activity and involves applying each artifact to solve a real-life problem as proof of concept to demonstrate its feasibility. In this activity, the applicability and effectiveness of each artifact were examined through scenarios within the SHR. The demonstration step also provided insights into how well the artifacts aligned with their intended purposes, supporting later evaluations.

Activity 5: Artifact evaluation. The artifact evaluation activity aims to determine how effectively each proposed artifact can solve the problem that motivated the research. As mentioned, this activity is closely related to the artifact demonstration activity. In Papers I-VIII, all created artifacts were evaluated and analyzed through experimental studies conducted in the SHR based on the considered policies or perspectives. The artifact evaluation in each paper is as follows:

- Paper II: We conducted a scenario study in the SHR to analyze the performance of the proposed optimization model by generating three MSU scenarios (including 1–3 MSUs) and used exhaustive search to solve the model.
- Paper V: We applied the MILP model for the MSU placement problem to the Blekinge and Kronoberg counties of Sweden to illustrate the functionality of our approach. We solved the model using an optimization solver to identify the optimal placements for different numbers of MSUs.
- Paper I: We evaluated the proposed macro-simulation model using two generated MSU scenarios and compared the experimental results with a baseline scenario (that is, using only RAs).
- Paper III: In a scenario study, we analyzed the outcome of the different stroke transport policies and their effectiveness. In addition, we performed a sensitivity analysis to evaluate the performance of the proposed DES model.
- Paper IV: We evaluated the framework in a scenario study by using it to construct a DES model to simulate EMS activities related to acute stroke, which was applied to the SHR.

- Paper VI: We used the extended framework to create an EMS simulation model for stroke patients, which we applied in a scenario study to the SHR. This allowed us to evaluate the effectiveness of using the co-dispatch policy for different types of stroke.
- Paper VII: We applied the proposed approach to the SHR to evaluate the potential benefits of using dynamic travel time calculations in the simulation of prehospital stroke care. We also compared the running time of our approach, a locally-deployed routing engine, with an online service.
- Paper VIII: In a scenario study, we applied the proposed approach to Skåne County, Sweden, to explore the potential impacts of using ML for RA travel time estimations in EMS simulations. We compared the estimated travel times for different travel activities both with and without using the ML-based approach in our simulation model.

The experimental results demonstrated the effectiveness of all proposed artifacts in order to, for example, identify optimal MSU locations to provide equitable and efficient access to stroke treatment, evaluating different stroke transport policies and MSU allocations, simplifying the construction of DES models for analyzing EMS policies, and enabling more realistic simulations of EMS scenarios and policies. To validate the experimental results, we compared them with baseline and earlier models and performed sensitivity analyses to assess robustness under varying conditions. Furthermore, the use of synthetic stroke data in this research served as a practical alternative that enabled us to conduct multiple simulations and model evaluations in the absence of detailed real-world data. Further details on the experimental results are provided in Chapter 3 Contributions.

Activity 6: Communication. In this step, the proposed artifacts, the acquired knowledge, and the research findings were communicated through presentations at sessions with stakeholders in the SHR and at scientific conferences. Additionally, the work was disseminated through publications in a scientific journal as well as conference proceedings within the research field. The communication step is essential for conveying the problem’s significance, the artifact as a proposed solution, and the artifact’s utility and effectiveness to both researchers and practitioners as well as other interested parties.

2.1. Data Collection

In the design science process, we used quantitative data to evaluate the proposed artifacts. We collected demographic data from Statistics Sweden [59], age-based stroke data from Sweden's Southern Regional Healthcare Committee [60], and ambulance spatiotemporal data from SOS alarm [61]. In addition, we divided the study region, that is, SHR, into a disjoint set of 1×1 km subregions to enable a detailed analysis of demographic variations and the expected times to treatment within the region. For simplicity, we assumed that all inhabitants/patients in a subregion are located at its center, meaning all transports to and from a subregion are considered to and from its center.

The demographic data included the number of inhabitants for each age group and each subregion under consideration in 2018. The stroke data provided hour-based statistics on reported stroke cases for 21 age groups in each municipality within the SHR in 2018, including the statistical distributions of the time, location, and age of the stroke patients. The ambulance spatiotemporal data (SOS Alarm) encompasses records of RA events from 2020 and 2021 for Skåne County, Sweden, describing RA transport activities corresponding to either emergency callouts or planned patient relocations. The dataset contained detailed information about each RA transport activity, such as the date and time the event was created, the assigned ambulance, the patient's location (given as the municipality for confidentiality), the ambulance's travel time to the patient's location, and the transport time from the patient's location to the hospital. Each event is assigned a priority level ranging from one (highest priority) to seven (lowest priority).

We used demographic data and stroke data as input for the proposed models in Paper I, Paper II, and Paper V. Moreover, we utilized both the stroke data and demographic data in a synthetic population generation model to create a synthetic stroke population, which was then used as input for the DES models in Paper III, Paper IV, Paper VI, and Paper VII. We also used ambulance spatiotemporal data (SOS Alarm) and synthetic stroke population in Paper VIII.

2.2. Research Method Discussion

One of the main goals of design science is to generate new knowledge that can assist decision-makers in designing solutions for problems within the area of concern. Therefore, applying the presented artifacts has the potential to support decision-makers in the domain to improve prehospital stroke care and the treatment of stroke patients, ultimately reducing human suffering and lowering the social costs associated with stroke.

In this thesis, we employed design science as the sole research method to explore, analyze, and address the proposed artifacts and research questions. However, it is important to note that this type of problem has also been studied within the field of operational research, which applies analytical methods to support decision-making, particularly in complex areas such as healthcare management.

3. CONTRIBUTIONS

This chapter outlines the research contributions of the publications included in this thesis in relation to the considered research questions. The contributions are organized and presented in a manner that corresponds to the research questions, beginning with RQ1. The connection between the research questions and publications is shown in Table 3.1.

Table 3.1. The relationship between the research questions and the papers included in this thesis.

RQ	Papers
RQ1	Papers II and V
RQ2	Papers I and III
RQ3	Papers IV and VI
RQ4	Papers VII and VIII

RQ1. How can optimization be used to determine the optimal placement of MSUs to provide equal and efficient access to stroke treatment?

Strategically placing MSUs in a region is expected to reduce the time to diagnosis and treatment for most inhabitants, a critical factor in improving patient outcomes. Mathematical optimization can be used to identify the best MSU locations based on the defined constraints, that is, the number of available EVs and the choice of how to make the trade-off between efficiency and equity. The optimal placement of an MSU in terms of either efficiency or equity is the focus of related studies in prehospital stroke care. However, no previous study proposes an optimization model to make a trade-off between efficiency and equity for the optimal placement of MSUs in a region. RQ1 is addressed in Papers II and V.

In Paper II, we proposed an optimization model for the MSU placement problem—in which the constraints are not mathematically formulated—and used an exhaustive search algorithm to solve it, aiming to achieve a trade-off between the efficiency and equity perspectives when allocating MSUs in a geographic region. We established a trade-off function for the MSU placement problem based on the concepts of weighted average time to treatment (WATT) and range, as presented in Equation 3.1. WATT was our measure of efficiency ($F_{efficiency}$), computed by summing each subregion's expected time to treatment multiplied by the fraction of the number of stroke incidents that is expected to occur in that subregion. Therefore, WATT effectively averages the time to treatment across all subregions, giving more influence to those with a higher expected number of stroke incidents.

The range was our equity measure (G_{equity}), which refers to the time difference between the expected time to treatment for inhabitants at different geographic locations. The aim of using the range is to quantify the difference between the maximum and minimum expected times to treatment among all subregions in the study region, thereby capturing how equitably the region is served; a smaller range implies greater equity in access to timely stroke care.

We used the trade-off function as the objective function in the optimization problem, where $W = 0.5$. In equation 3.1, W is a controlling weight, assigned by the user, which is used to determine the impact of each of efficiency and equity on the allocation of MSUs. In a special case, when $W = 0$ or $W = 1$ in Equation 3.1, only efficiency or equity is considered, respectively. The use of the weight W in the trade-off function enables supporting the public health authorities' priorities regarding efficiency and equity when deciding where to place MSUs.

$$(1 - W)\left(F_{efficiency}\right) + W\left(G_{equity}\right) \tag{3.1}$$

In Paper II, we solved the MSU placement problem using an exhaustive search algorithm. However, the use of MILP can be beneficial in representing the essential characteristics of the MSU placement problem. One key advantage of using advanced solution finding methods for MILP formulations over exhaustive search is scalability, which enables handling larger-scale problems that would be computationally infeasible with exhaustive search. In addition, solving MILP formulations using advanced optimization solvers and methods can potentially reduce computational time for complex problems by finding exact solutions without evaluating every possible combination. However, no prior research has directly addressed the mathematical formulation of the MSU placement problem. Therefore, to further answer RQ1, in Paper V, we introduced an MILP model to mathematically represent the MSU placement problem, where the model's objective function was the mentioned trade-off function. This objective function allows healthcare planners to balance efficiency and equity in MSU placement, making it adaptable to various regional priorities.

In summary, the contributions of Paper II and V in addressing RQ1 are as follows:

- A trade-off function to balance the efficiency and equity perspectives for optimal placement of different numbers of MSUs across a large region, encompassing both urban and rural areas.
- An MILP model to represent the MSU placement problem, designed to optimize MSU allocation by incorporating the trade-off in the objective function.

RQ2. How can simulation be used to evaluate stroke transport policies?

In prehospital stroke care, prior studies focus mainly on developing simulation models to assess stroke transport policies, including only RAs. To the best of our knowledge, no previous study explicitly uses simulation modeling to study stroke transport policies related to MSUs and to study the effects of different MSU placements. RQ2 is addressed in Papers I and III.

In Paper I, we extended the work of Dahllöf et al. [8], who study the potential benefits of using an MSU in Skåne County, which is part of the SHR. We presented a macro-level simulation model using an average time to treatment estimation model to analyze the potential benefits of placing MSUs in the SHR. The contribution of Paper I was twofold: i) a model to estimate the expected time to treatment for the entire population and different parts of a geographic region, and ii) the results of a scenario study where the model was applied within the SHR.

The proposed approach in Paper I was a macro-level simulation model, which provided an overview of the problem and enabled us to analyze the outcomes on a large scale. However, the use of a micro-level simulation model can be beneficial when it comes to studying the behavior of individual patients and individual EVs over time, leading to more realistic modeling. It also enables one to incorporate stochasticity into the model, such as the location and time of stroke incidents, and the effects of simultaneous stroke incidents, which sometimes places the demand on an MSU to be at two places at the same time.

In Paper III, we contributed with a microscopic simulation model built according to the DES paradigm to evaluate the potential benefits of MSUs in the SHR. In the prehospital stroke care chain, almost all activities and events occur as a sequence of discrete events; hence, we consider DES to be a preferable choice to model such systems. The proposed DES model in Paper III is able to simulate the main activities and decisions involved in the prehospital stroke care chain and uses a synthetic population of stroke patients as input.

Due to the privacy restrictions and limitations in our stroke data (for example, patients' exact locations were unknown, and patient data did not include all the required attributes), we utilized a synthetic stroke population generation model to create different sets of synthetic stroke populations, which were then used as input for the DES model. This approach preserved privacy and enabled the study of different scenarios and the validation of the simulation model through sensitivity analyses. To generate the synthetic stroke population, we used real aggregated stroke data [60] and demographic data [59]. Each patient in the synthetic stroke population had different attributes, including incident time (hour, day), location, and age. In the synthetic population generation model, we assumed that the occurrence of stroke incidents follows a Poisson distribution. In particular, the Poisson distribution models the expected number of stroke incidents within each time period (that is, the hours of the day), and the time interval between consecutive incidents is modeled using an exponential distribution. For each stroke incident, patient attributes such as location and age were sampled based on aggregated statistics.

In summary, the contributions of Papers I and III in addressing RQ2 are as follows:

- A macro-level simulation model is introduced using an average time to treatment estimation model to explore the potential benefits of deploying MSUs across a large region.
- A micro-level DES model is used to assess various stroke transport policies, including the use of MSUs, enabling the study of individual entities over time and incorporating stochastic elements into the model.

RQ3. What is an appropriate modeling framework for constructing simulation models to analyze real-world EMS operations?

A modeling framework is a useful tool for constructing simulation models. The key advantage of a modeling framework is its simplicity in generating specific models. However, no prior study explicitly contributes a modeling framework for constructing DES models for the analysis of EMS policies. RQ3 is addressed in Papers IV and VI.

To address RQ3, we proposed a modeling framework in Paper IV in order to construct DES models to analyze EMS policies, leading to simplifying the construction of DES models in EMS. The framework was established on the idea of modeling EMS care chains using standard flow charts consisting of connected activities and decisions. The main activities identified in most EMS care chains are included in the proposed framework, regardless of the diagnosis of the patient. As presented in Figure 3.1, the framework consists of two main phases: model construction and model execution. In the model construction phase, various inputs, such as geographic data, patient data, ambulance data, hospital data, and a care chain specification, are utilized to create a DES model for a specific EMS scenario. In the execution phase, the constructed model is run with the input data, and the specified care chain is applied to each simulated patient. An example of a care chain of EMS activities for a suspected stroke patient, including the decision-making activities, is illustrated in Figure 2 in Paper IV.

The framework proposed in Paper IV only supported a single dispatch policy, and in the experiments, we assumed that the synthetic population only consisted of ischemic stroke patients requiring thrombolysis. Similarly, previous research focuses on different aspects of prehospital stroke patient transportation, but has not conducted a simulation-based evaluation of co-dispatching in prehospital stroke care. To address this gap, in Paper VI, we extended the existing simulation framework by incorporating various dispatch policies, including co-dispatch, to evaluate the potential benefits of co-dispatching RAs and MSUs in prehospital stroke care. In addition, we used a synthetic stroke population with different stroke types and treatment requirements as input for the framework. In Figures 2

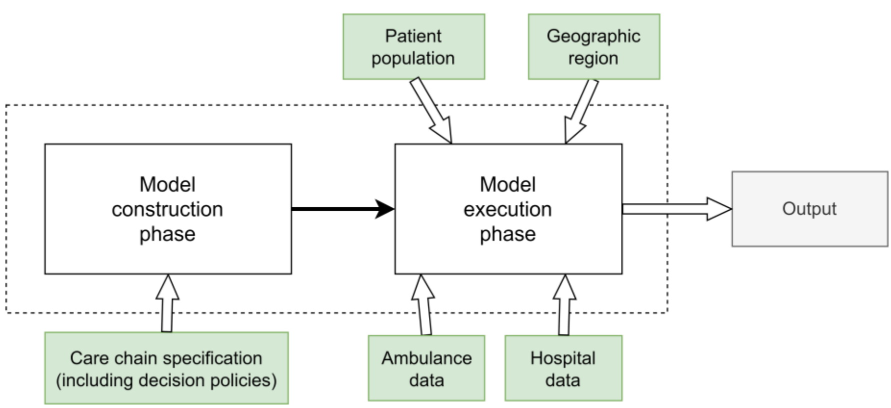


Figure 3.1. Overview of the modeling framework, from Paper IV, where the provided input data is used to construct and run the generated simulation model.

to 4 in Paper VI, we presented the proposed EMS care chain of activities, including the co-dispatch policy, for a suspected stroke patient. In our co-dispatch policy, the patient's location served as the meeting point for the two EVs. The extended framework could handle either single dispatch or co-dispatch for each incident, as well as the interaction between EVs in co-dispatching, depending on which EV (RA or MSU) arrives first.

In summary, the contributions of Papers IV and VI in addressing RQ3 are as follows:

- A simulation modeling framework that simplifies the construction of DES models for analyzing EMS policies. It offers high flexibility to adapt to different scenarios and EMS care chains—including complex cases like acute stroke.
- An evaluation of the framework by applying it to construct a DES model for stroke patients.
- The incorporation of a co-dispatch policy into the framework to explore the potential benefits of co-dispatching for stroke patients requiring different treatments.

RQ 4. How does integrating dynamic travel time calculation and ML-based travel time estimation into the EMS simulation modeling framework impact RA travel time estimations?

In the original framework presented in Paper IV, we used pre-calculated travel times for EVs generated through Open Street Map (OSM) via the ORS toolbox in Quantum Geographic Information System (QGIS). Generating pre-calculated travel times often requires a large number of individual requests to a routing engine, which can lead to additional steps like partitioning large datasets and manually merging multiple outputs. This process of preparing and collecting such a large amount of travel data can be complex, time-consuming and prone to errors, thereby limiting the flexibility of the simulation models, especially when adapting the framework to different regions. In addition, the use of pre-calculated travel times restricts EV dispatch to a fixed set of predefined locations, preventing dynamic dispatch from other locations within the region. While a dynamic approach for travel time calculations could overcome these limitations, the impact of integrating dynamic travel times into EMS simulations, particularly for prehospital stroke care, remains largely unexplored. RQ4 is addressed in Papers VII and VIII.

To address these limitations and RQ4, in Paper VII, we integrated a routing engine, that is, OSRM-CH, locally into the framework, enabling dynamic travel time calculations. OSRM-CH outperforms other routing engines based on different criteria, such as those based on licensing flexibility, matrix search capability, request limitations, and usability for local deployment, as highlighted by Juninger and Narvell [51]. In Figure 2 of Paper VII, we illustrated how the modeling framework and the dynamic travel time calculation module interact to calculate travel times during the simulation. An instance of the OSRM-CH routing engine is configured and run locally on the Docker container. During the simulation, when a request for calculating the travel time between two locations is sent to the module, a response containing the corresponding travel time and distance is generated and returned.

Accurate travel time estimations of RAs are essential for making EMS simulations more reliable and effective. Although using a routing engine for dynamic travel time calculations in the framework is beneficial, it typically does not accurately estimate RA travel times. The travel times calculated with OSRM-CH represent average travel times for passenger cars, which likely differ from actual RA travel times and may not fully reflect the conditions of EMS operations during simulations. To partially address this issue, in our earlier papers, we used a driving adjustment factor to adjust travel times calculated by routing engines such as ORS or OSRM-CH. However, some studies suggest that ML approaches can further improve RA travel time estimations [53-55]. ML approaches can be used to correct the travel times generated by routing engines to better approximate RA travel times. Despite this potential, the benefits of using ML-based approaches for estimating RA travel times in EMS simulations remain unexplored.

To answer RQ4, in Paper VIII, we integrated an ML-based RA travel time estimation module into the EMS simulation framework to enable more realistic travel time estimations for RAs. The ML module, trained on spatiotemporal ambulance data (SOS Alarm), adjusts travel time calculations based on features such as priority level, hour of the day, day of the week, distance, and dynamically calculated travel times. In the module, an ANN model was selected and trained using these features. As shown in Figure 2 of Paper VIII, during the simulation, the travel time for the two requested locations is first calculated using the dynamic travel time calculation module. This result, along with other mentioned features, is then used as input for the trained ANN model in the ML-based RA travel time estimation module to correct the RA travel time.

In summary, the contributions of Papers VII and VIII in addressing RQ4 are as follows:

- Dynamic travel time calculation functionality is incorporated into the framework to enhance flexibility and adaptability, allowing simulations to better reflect real-world scenarios by removing limitations of pre-calculated travel data.
- The integration of an ML-based travel time estimation module into the framework to improve accuracy in estimating RA travel times and to assess its impact in EMS simulation.

3.1 Scenario Study Discussion

To further emphasize the practical implications of our contributions, this section describes the scenario studies and presents the key findings from the publications included in this thesis, addressing the research questions.

RQ1. How can optimization be used to determine the optimal placement of MSUs to provide equal and efficient access to stroke treatment?

In a scenario study, we evaluated the proposed optimization model in Paper II by comparing the current situation, including only RA, in the SHR with three generated MSU scenarios, each including one, two, and three MSUs.

The computational results demonstrated that the suggested optimization model could create a balance between efficiency and equity in MSU allocations by reducing both WATT and range. For example, the results of the trade-off when three MSUs are placed optimally showed that WATT decreased by 16.8 minutes, and the range was reduced by 18.6 minutes compared to the baseline. The results also indicated that after placing one, two, and three MSUs using the trade-off, the time to treatment is expected to decrease for 29%, 59%, and 75% of the total population, respectively. In particular, placing 3 MSUs in the SHR is expected to lead to an improvement by approximately a factor of 14 in the share of persons who get treatment within an hour.

In Paper V, to validate the MILP model's functionality and performance on a more manageable scale, we conducted a scenario study by applying the model to two smaller counties within the SHR, Blekinge and Kronoberg. We solved the model using the barrier and simplex algorithms provided by the Gurobi solver to determine the optimal placements for various numbers of MSUs. We compared the performance of these algorithms in terms of execution time and scalability for large problem instances.

The experimental findings, supported by the results from the exhaustive search algorithm, demonstrated the correctness of the proposed model. From the experiments, we explored the scalability of our optimization model and assessed the limits of using the Gurobi solver for the MSU placement problem, where the time required to find an optimal solution provided insights into the problem's computational complexity. In addition, the optimal MSU locations identified by the model were shown to reduce the expected time to treatment compared to the baseline. Notably, the results indicated that placing two MSUs in Blekinge and five in Kronoberg would likely provide access to treatment within an hour for all residents, thus achieving an important healthcare goal. Moreover, in the same scenario, compared to the baseline, the average time to treatment decreased from 1.61 hours to 0.97 hours in Blekinge and from 1.78 hours to 0.98 hours in Kronoberg.

RQ2. How can simulation be used to evaluate stroke transport policies?

We applied the proposed model in Paper I in a scenario study within the SHR. In particular, we compared the current situation with two extended MSU scenarios (denoted as MSU1 and MSU2), where we added three MSUs to each. The experimental results showed that the use of MSUs is expected to decrease the expected time to treatment for the whole region, for each municipality, and for each subregion covering SHR, and how the population is expected to benefit from MSUs. Compared to the baseline scenario, the average time to treatment dropped from 1.62 h to 1.28 h and 1.36 h for MSU1 and MSU2, respectively.

In addition, in MSU1, MSUs were placed in a way that potentially improved the situation for a larger share of the population of SHR (efficiency perspective). For example, the time to treatment is expected to decrease by 82% and 48% of the total population of SHR for the MSU1 and MSU2 scenarios compared to the baseline scenario, respectively. In contrast, in MSU2, MSUs were placed to potentially improve the situation for a larger share of those patients who currently have the longest time to treatment (equity perspective). Deploying MSU1 reduced the average time to treatment to less than 2 hours in 87.4% of subregions and for 93.7% of people, compared to the baseline. Deploying MSU2 reduced the average time to treatment to less than 2 hours in 98.8% of subregions and for 99.4% of people, compared to the baseline. This indicates that MSU2 had a higher effect in reducing the time to treatment for those who currently have longer time to treatment (for example, more than 2 hours).

To observe the effectiveness of the proposed DES model in Paper III, we compared the time to treatment, the number of MSU dispatches, and the average dispatching distance of MSUs for the different scenarios (similar to Paper I: MSU1 and MSU2) using the simulation outputs. We also conducted a sensitivity analysis to further evaluate the presented DES model, that is, varying the number of MSUs in each site and the number of patients in the synthetic population.

The simulation results showed that the use of MSUs is expected to reduce the time to treatment in the SHR and to help more stroke patients get treatment earlier. Another finding was that when MSUs are located in or near densely populated areas, the number of MSU dispatches is expected to considerably increase as more patients benefit from the MSUs. In addition, MSUs positioned in populated areas could reach farther distances than those placed in rural areas. For example, the average dispatching distance is 45.38 km and 40.54 km for the MSUs in MSU1 and MSU2, respectively.

Comparing the average time to treatment between the DES model and macro-level simulation model indicated that the DES model, which accounts for limited EV availability and simultaneous stroke incidents, has a slightly longer time to treatment for all scenarios. In addition, the sensitivity analysis showed that placing 10 MSUs per site in the DES model produced results similar to the macro-level model, as the increased MSU availability eliminates resource limitations, which is the case in the macro-level model. These findings demonstrate that the DES paradigm provides a more realistic representation than the macro-level simulation model (Paper I).

Comparing the results of the optimization papers (Papers II and V) with the simulation papers (Papers I and III) shows that the optimal placement of MSUs would lead to a shorter time to treatment and increase the share of inhabitants/patients expected to receive treatment earlier. In contrast, in the simulation papers, MSU locations were selected based on a visual analysis of the results of the baseline scenario rather than an optimization-based approach. In Paper I, the MSUs in the MSU1 and MSU2 scenarios were placed to address efficiency and equity, respectively.

Comparing the results of the baseline with the MSU1 scenario in Paper I and the optimal placement of three MSUs in Paper II (focusing on efficiency) shows that WATT, our measure of efficiency, decreased by 9 minutes in Paper I and 17.4 minutes in Paper II. Similarly, when comparing the share of inhabitants expected to receive treatment within one hour, the increase was by a factor of 3.4 in Paper I and 14.6 in Paper II. In addition, comparing the baseline results with the MSU2 scenario in Paper I and the optimal placement of three MSUs in Paper

II (considering equity) shows that the range, our measure of equity, decreased by 12.24 minutes in Paper I and 20.4 minutes in Paper II. These comparisons emphasize the importance of using optimization for the strategic placement of MSUs in a region.

RQ3. What is an appropriate modeling framework for constructing simulation models to analyze real-world EMS operations?

In a scenario study, we utilized the framework proposed in Paper IV to build a model for simulating EMS activities associated with acute stroke, which has a more complex care chain than the conventional EMS chain. The main reason is that the use of MSUs in the prehospital stroke care chain allows one to perform diagnosis and treatment at the patient's location, and incorporating these specifications into the prehospital stroke care chain introduces multiple decisions that need to be taken.

The simulation results showed that the use of MSUs, compared to the baseline scenario, enables to provide better service for stroke patients in the SHR. In particular, when the decision for the choice of an EV is made using each of the *time to diagnosis* and *time to treatment* policies. The improvement is more considerable when the hospital selection policy is *direct to special clinic*. In addition, the simulation results presented in Paper IV are similar to those in Paper III, demonstrating that the constructed simulation model for the stroke care chain functions as intended and that the general care chain activities included in the framework were sufficient, at least for modeling the relatively complex case of acute stroke.

We applied the extended framework in Paper VI in a scenario study to SHR to evaluate the effectiveness of using the co-dispatch policy for different types of stroke. In the experiments, we considered various decision policies, including EV selection, hospital selection, dispatch type selection, and transferring to hospital policies.

The simulation results demonstrated that co-dispatch generally outperforms single dispatch for stroke patients receiving different types of treatment by reducing the time to both diagnosis and treatment, particularly when three MSUs are placed in the SHR. For example, using co-dispatching with three MSUs is expected to reduce the average time to treatment by 11, 25, 32, and 25 minutes for ischemic stroke patients receiving no treatment, thrombolysis, thrombectomy, and both thrombolysis and thrombectomy, respectively, compared to the baseline. In addition, for hemorrhagic cases, the average time to initiate blood pressure-lowering therapy is expected to decrease by approximately 13 minutes. Under the same scenario, the co-dispatch policy is applied in 61.40% of all dispatches.

However, for patients requiring thrombectomy, co-dispatching may result in a longer time to diagnosis than single dispatch, as the CT scanner in the MSU cannot fully determine the specific treatment(s) needed for suspected ischemic stroke patients [62]. Consequently, in the simulation, we assumed that ischemic stroke patients potentially requiring thrombectomy should be transported to a hospital for further diagnosis, even if they were initially diagnosed inside the MSU. To address this limitation, advancements in the CT scanner technology currently installed in MSUs are needed to enable them to differentiate the specific type of treatment an ischemic stroke patient requires [63].

RQ 4. How does integrating dynamic travel time calculation and ML-based travel time estimation into the EMS simulation modeling framework impact RA travel time estimations?

In Paper VII, we evaluated the extended framework by constructing a simulation model for prehospital stroke care, which we applied to the SHR. In the experiments, we compared dynamic travel time calculations with the pre-calculated data approach, evaluating both in terms of average time to treatment, running time, and data loading requirements during simulation.

In the experiments, the routing engine used for the dynamic approach was OSRM-CH and for the pre-calculated approach was the ORS toolbox in QGIS. Although we used different routing engines to calculate travel data, the times to diagnosis and treatment were approximately equal across all scenarios for both dynamic and pre-calculated travel times. This demonstrates that the proposed approach for integrating the dynamic travel time calculations of EVs into the framework functions as intended.

The dynamic calculation approach eliminates the need for extensive data loading required by the pre-calculated approach. For example, loading the pre-calculated travel data for SHR into the framework before starting the simulation took about 22 minutes on average. In larger regions, calculating and loading travel data would be even more time-consuming, making the pre-calculated approach less adaptable to new regions.

The results showed that the total simulation time, summing travel data loading time and run time, for the pre-calculated approach was about 48 minutes on average. In contrast, the total simulation time for the dynamic approach, which does not require any travel data loading, was 67 minutes on average. Although the dynamic approach requires more running time per simulation, its impact on the flexibility and adaptability of EMS simulation models makes it a valuable addition to the framework for dynamic travel time calculations.

We also compared the performance of an online service with our proposed dynamic approach by randomly selecting 500 pairs of coordinates between patient and ambulance locations in SHR. The results for processing and calculating 500 requests indicated the online service was 44 times slower than the local routing engine integrated into our framework.

In Paper VIII, we evaluated the extended framework by constructing a simulation model for prehospital stroke care in Skåne County, Sweden. We compared emergency response metrics with and without using the ML-based RA travel time module. The emergency response metrics included time to diagnosis, time to treatment, and time for three ambulance transport activities: ambulance-to-patient, patient-to-hospital, and hospital-to-hospital.

The results showed that using the ML-based approach for RA travel time estimations resulted in differences across various emergency response metrics compared to using only dynamic travel times. In addition, the distribution of estimated travel times for transport activities across different travel distances (see Figure 4 in Paper VIII) highlighted the differences between estimations with and without the ML-based approach, particularly for longer trips (over 20 km). This suggests that RA driving speeds differ from those of passenger vehicles and highlights the potential impact of ML-based estimations on EMS simulation outcomes. However, due to data limitations (for example, the ambulance spatiotemporal data also included non-stroke and non-emergency incidents, and the patient's exact location was replaced by the municipality name), the results did not demonstrate a clear improvement in the accuracy of estimated ambulance travel times.

We also compared the execution time of the simulation model applied to Skåne County with and without using the ML-based approach. Our analysis revealed that incorporating the ML-based estimation module increased execution time by 5.79 minutes. However, the impact on simulation accuracy justifies this additional computational cost and supports using ML approaches for more accurate RA travel time estimation.

4. CONCLUSIONS AND FUTURE WORK

4.1. Conclusions

The main aim of this thesis was to explore how simulation and optimization can improve the analysis and planning of prehospital stroke care. The focus was on studying the application of MSUs in prehospital stroke care using these approaches. In particular, we showed that optimization can be useful for identifying the optimal MSU locations considering, for example, the number of available EVs and the choice of how to trade-off between efficiency and equity. Simulation enabled us to evaluate different stroke transport and dispatch policies, including different MSU placements, for stroke patients with different treatment needs. The aim of the thesis was achieved by addressing four specific research questions, each corresponding to a specific objective of the thesis.

The first objective (RQ1) was to build an optimization model to identify optimal MSU locations that create a trade-off between the efficiency and equity perspectives. To address the first objective, in Paper II, we formulated a trade-off function between these two perspectives for the MSU placement problem, which was solved using an exhaustive search algorithm. In Paper V, we further expanded on the optimal MSU placements by presenting an MILP model. The results demonstrated that the identified optimal MSU locations in the SHR and the proposed trade-off function could provide equitable and efficient access to stroke treatment for most residents. Additionally, comparing the results from solving the MILP model with a standard solver and those from Paper II—where our optimization model was solved using an exhaustive search—confirmed the correctness of the proposed MILP model. However, the results also revealed the computational challenges in scaling the MILP model for larger regions like the SHR, particularly when using standard optimization solvers. This highlights the need for scalable optimization approaches to support EMS planning in complex regions.

The second objective (RQ2) involved using simulation to assess stroke transport policies, especially those incorporating MSUs. To achieve this objective, in Paper I, we proposed a macro-level simulation model to explore the potential benefits of deploying MSUs on a larger scale within the SHR. Furthermore, in Paper III, we presented a DES model for a more granular assessment of stroke transport policies within the SHR, including MSU deployment. The results demonstrated that MSUs could increase the proportion of the population receiving timely treatment and reduce the time to treatment for the entire region. From the results, we concluded that different MSU locations benefit inhabitants living in different parts of the SHR in different ways. In particular, placing MSUs in densely populated areas significantly increases the number of MSU dispatches, further reducing the expected time to treatment. Based on the simulation results and conducted sensitivity analysis, we observed the advantages of micro-level simulations, like DES, for analyzing stroke care in greater detail and concluded that the developed DES paradigm could provide more realistic results than the macro-level modeling approach.

The third objective (RQ3) was to create a generalized framework that builds simulation models for EMS policy analysis. To meet this objective, in Papers IV and VI, we presented a modeling framework that provides a structured, reusable approach to simplify the construction of DES models for EMS policy analysis. The framework incorporates different decision policies, including EV selection, dispatch type (single and co-dispatch) selection, and hospital selection, allowing for the evaluation of stroke transport policies across different stroke types. The proposed framework is flexible and can be adapted to any region, allowing end-users to customize it with local EMS resources, geographic needs, and appropriate inputs. The experimental results for evaluating the functionality of the framework indicated that the framework's components were adequate for modeling the complexities of the EMS care chain, such as in acute stroke care. The simulation results also demonstrated the advantages of co-dispatch over single dispatch by showing reductions in time to both diagnosis and treatment for stroke patients receiving various types of care. Notably, the use of MSU-dispatch or co-dispatch policies significantly reduced the time to treatment for patients requiring thrombectomy by minimizing inter-hospital transfers. The results also underscored the value of deploying MSUs, whether in single- or co-dispatch policies, for hemorrhagic stroke cases.

The fourth objective (RQ4) was to integrate dynamic travel time calculations (Paper VII) and ML-based travel time estimation (Paper VIII) into the framework to assess their impact on RA travel time estimations in EMS simulations. The results highlighted that using dynamic calculations could enhance the flexibility and adaptability of EMS simulations, eliminate the need for pre-calculated travel times, and reduce setup complexity, although it increased the simulation runtime. In addition, a comparison between local-based and online services for dynamic travel time calculations demonstrated that the local implementation used in our approach is faster and less restricted regarding the number of requests compared to online services. Furthermore, comparing the results of using the ML-based approach and dynamic approach for RA travel estimations demonstrated that the ML-based approach could provide more accurate RA travel time estimations, potentially improving the reliability of EMS simulations. Additionally, incorporating the ML module into the framework introduced only slight computational complexity, making it a feasible and beneficial addition to EMS simulation models.

In summary, the findings of this thesis demonstrate the potential of optimization and simulation to support the analysis and planning of stroke transport policies and to improve decision-making in prehospital stroke care. Optimized MSU placements and co-dispatch policies were shown to potentially reduce the time to diagnosis and treatment for stroke patients, while the integration of dynamic and ML-based travel time estimations could enhance the flexibility and accuracy of simulations. Delivering faster stroke care can potentially reduce stroke-related disabilities, leading to better patient outcomes and quality of life. Additionally, faster access to stroke treatments can reduce the long-term social and healthcare costs associated with stroke, thereby justifying the investment in MSU deployment within a region. These findings underscore the value of these approaches in guiding EMS policy design, ultimately contributing to improved stroke care delivery and reduced social impacts of stroke. The results of this thesis could provide practical insights for public health authorities when making decisions in prehospital stroke care.

4.2. Future work

Based on the findings presented in this thesis, three potential directions for future work are outlined below.

4.2.1. Developing Optimization Approaches for MSU Placement and Co-dispatching

In line with our contributions to the use of optimization in addressing the MSU placement problem, there are relevant optimization-based directions for further exploration. For example, the results in Paper V demonstrated that solving the proposed MILP model for large problem instances, such as those involving the SHR, is computationally expensive, especially when using standard optimization solvers. A potential future direction is to explore the use of heuristics to solve the proposed MILP model for large problem instances within a reasonable time frame.

In Paper VI, we assumed the patient’s location as the meeting point when co-dispatching an RA and an MSU for an incident. The results showed that this choice led to considerable waiting times for the EVs at the patient’s location. The extended framework now supports dynamic travel time calculations, enabling dynamic calculation of travel times between any two locations within the study region. This functionality allows to dynamically determine optimal meeting points based on the current locations of the patient, RA, and MSU, potentially reducing waiting times and enhancing service efficiency. For future work, we propose an MILP model to identify optimal meeting points for co-dispatching, aiming to minimize time to diagnosis and treatment. Additionally, we plan to expand the framework to integrate this optimization model.

In our experiments, we assumed ambulance sites as potential locations for placing MSUs and conducted experiments based on this assumption using the proposed optimization and simulation models. In future work, it would also be relevant to consider hospital locations as potential sites for placing MSUs. This approach could help reidentify the optimal MSU locations in the study region and use these identified locations in simulations to assess stroke transport policies. The results could then be compared to those obtained from the current assumption, providing further insights into the implications of MSU placement strategies.

4.2.2. Enhancing Simulation Models through Data Integration and Validation

As this thesis contributed to simulation-based approaches for analyzing stroke transport policies, future work could explore enhancements in simulation modeling and data integration. In Paper VIII, we integrated ML-based travel time estimation functionality into the framework to enhance simulation modeling by refining RA travel time estimations. The experimental results showed that this functionality impacted simulation outcomes. In addition, we found that the effectiveness of the ML-based estimation module depends on the quality and quantity of training data. One limitation of the current RA spatiotemporal dataset (SOS Alarm) used in Paper VIII was that it encompassed all types of RA transport activities, including non-emergency ones, rather than solely focusing on stroke or high-priority cases. Furthermore, the patients' locations were reported by municipality names rather than precise coordinates to protect the patients' confidentiality. This generalization may limit the ML-based module's effectiveness in estimating travel times for emergency cases. With more granular data—particularly data that includes the patients' exact locations and differentiates between types of emergencies, priority levels, and contextual factors like weather conditions—the performance of the ML-based module could be further refined. For future work, we plan to submit an ethical application to access more detailed RA data. In addition to aiming to improve RA travel time estimations, it is also relevant to explore how better estimations of MSU travel times can be achieved in simulation.

The stochasticity in our simulation models was primarily reflected in the patient population, where each patient has different random attributes and was randomly distributed across the study region. Additional sources of stochasticity were incorporated by considering that an EV assigned to an operation remains unavailable until it is dismissed. This typically occurs when the patient is left at their location, transported to a hospital, or transferred to another EV. We have analyzed these uncertainties through various scenarios and sensitivity analyses. As a future direction, the stochasticity in the proposed simulation models could be enhanced by integrating additional real-world factors. In addition, incorporating details such as road network characteristics, EV operating hours, and the number of RAs at each station would enable the models to better reflect real-world EMS operations.

Due to limitations in the available stroke data (for example, patients' exact locations were unknown), we used the synthetic stroke population as input for the proposed simulation models. This synthetic data approach allowed us to explore diverse scenarios and perform sensitivity analyses while preserving privacy and approximating real-world conditions. However, real-life validation using detailed datasets from the study region, such as all stroke cases and ambulance positions over multiple years, remains a future step to further validate the proposed models. Moreover, in the experiments, we mainly focused on the distribution of patients, their diagnoses, and preferred treatments, assuming that stroke affects all individuals to the same extent, even though this assumption may not reflect reality. Another possible future direction is to investigate the impact of all patient attributes on the simulation results, for example, examining how different MSU placements and stroke transport policies impact specific subgroups, such as individual age groups.

We validated the experimental results of the proposed models by comparing them with baseline and earlier models, and conducting sensitivity analyses to assess their robustness under varying conditions. A potential direction for future work is to evaluate the simulation outcomes, such as time to diagnosis and time to treatment, by comparing them with real-world data from the SHR. Such a comparison could provide deeper insights into the accuracy and reliability of the proposed simulation models and is an important future direction.

4.2.3. Health Economy Study: Evaluating the Cost-effectiveness of MSU Deployment and Co-dispatch Policy

In this thesis, we highlighted the time-oriented benefits of different MSU placements in the SHR and MSU-related stroke transport policies, including co-dispatching. Given the significant investment required for MSU deployment, it is essential to explore the health- and transport-economic implications of the use of MSU in the SHR. Some earlier studies review the economic aspects of stroke care and transportation, such as exploring the cost-effectiveness of optimizing the number and locations of special clinics in a wide region [64], operating MSUs [12, 14], and utilizing ambulance helicopters [65].

In future work, we aim to incorporate a cost function for MSUs into our optimization model to analyze the cost-effectiveness of MSU deployment compared to conventional stroke treatment costs, using available health economic data on MSU operations across a region. As a preliminary step, we plan to conduct a literature review to identify relevant measures to quantify costs for patients, RAs, and MSUs. In addition, although our experimental results demonstrated the advantages of co-dispatching in reducing the time to diagnosis and treatment for stroke patients, it is important to explore its cost-efficiency compared to single dispatch policy. Assessing the economic trade-offs of co-dispatching could further emphasize the benefits of this policy and provide a stronger justification for its practical implementation.

REFERENCES

[1] V. L. Feigin, M. Brainin, B. Norrving, S. Martins, R. L. Sacco, W. Hacke, M. Fisher, J. Pandian, and P. Lindsay, "World Stroke Organization (WSO): global stroke fact sheet 2022," *International Journal of Stroke,* vol. 17, no. 1, pp. 18-29, 2022.

[2] The Swedish Stroke Register, "Stroke Registrations (Annual Reports)," [Online]. Available: https://www.riksstroke.org/sve/forskning-statistik-och-verksamhetsutveckling/rapporter/arsrapporter/. [Accessed: Oct. 1, 2024].

[3] World Stroke Organization, “Facts and figures about stroke,” [Online]. Available: https://www.world-stroke.org/world-stroke-day-campaign/why-stroke-matters/learn-about-stroke/. [Accessed: Oct. 1, 2024].

[4] J. L. Saver, "Time is brain—quantified," *Stroke,* vol. 37, no. 1, pp. 263-266, 2006.

[5] T. Leng and Z.-G. Xiong, "Treatment for ischemic stroke: From thrombolysis to thrombectomy and remaining challenges," *Brain circulation,* vol. 5, no. 1, p. 8, 2019.

[6] M. Ebinger, B. Winter, M. Wendt, J. E. Weber, C. Waldschmidt, M. Rozanski, A. Kunz, P. Koch, P. A. Kellner, and D. Gierhake, "Effect of the use of ambulance-based thrombolysis on time to thrombolysis in acute ischemic stroke: a randomized clinical trial," *Jama,* vol. 311, no. 16, pp. 1622-1631, 2014.

[7] Statistics Sweden, “Folkmängd och befolkningsförändringar - Kvartal 1, 2021 - Kvartal 4 2023,” [Online]. Available: https://www.scb.se/hitta-statistik/statistik-efter-amne/befolkning/befolkningens-sammansattning/befolkningsstatistik#_Tabellerochdiagram/. [Accessed: Sep. 19, 2024].

[8] C. Patton, D. Sawicki, and J. Clark, *Basic methods of policy analysis and planning*. Routledge, 2015.

[9] K. Fassbender, J. C. Grotta, S. Walter, I. Q. Grunwald, A. Ragoschke-Schumm, and J. L. Saver, "Mobile stroke units for prehospital thrombolysis, triage, and beyond: benefits and challenges," *The Lancet Neurology,* vol. 16, no. 3, pp. 227-237, 2017.

[10] K. Fassbender, S. Walter, Y. Liu, F. Muehlhauser, A. Ragoschke, S. Kuehl, and O. Mielke, "“Mobile stroke unit” for hyperacute stroke treatment," *Stroke,* vol. 34, no. 6, pp. e44-e44, 2003.

[11] R. Cerejo, S. John, A. B. Buletko, A. Taqui, A. Itrat, N. Organek, S. M. Cho, L. Sheikhi, K. Uchino, and F. Briggs, "A mobile stroke treatment unit for field triage of patients for intraarterial revascularization therapy," *Journal of Neuroimaging,* vol. 25, no. 6, pp. 940-945, 2015.

[12] J. Kim, D. Easton, H. Zhao, S. Coote, G. Sookram, K. Smith, M. Stephenson, S. Bernard, M. W Parsons, and B. Yan, "Economic evaluation of the Melbourne mobile stroke unit," *International Journal of Stroke,* vol. 16, no. 4, pp. 466-475, 2021.

[13] H. Zhao, S. Coote, D. Easton, F. Langenberg, M. Stephenson, K. Smith, S. Bernard, D. A. Cadilhac, J. Kim, and C. F. Bladin, "Melbourne Mobile Stroke Unit and Reperfusion Therapy: Greater Clinical Impact of Thrombectomy Than Thrombolysis," *Stroke,* vol. 51, no. 3, pp. 922-930, 2020.

[14] U. H. Lund, A. Stoinska-Schneider, K. Larsen, K. G. Bache, and B. Robberstad, "Cost-effectiveness of mobile stroke unit care in Norway," *Stroke,* vol. 53, no. 10, pp. 3173-3181, 2022.

[15] MEYTEC GmbH, "Mobile Schlaganfallklinik mit Telemedizin-Ausstattung," [Online]. Available: https://meytec.de/stemo-msu/. [Accessed: Dec. 20, 2024].

[16] P. Hariharan, M. B. Tariq, J. C. Grotta, and A. L. Czap, "Mobile stroke units: current evidence and impact," *Current neurology and neuroscience reports,* vol. 22, no. 1, pp. 71-81, 2022.

[17] S. Mathur, S. Walter, I. Q. Grunwald, S. A. Helwig, M. Lesmeister, and K. Fassbender, "Improving prehospital stroke services in rural and underserved settings with mobile stroke units," *Frontiers in neurology,* vol. 10, p. 159, 2019.

[18] N. Glober, M. Supples, S. Persaud, D. Kim, M. Liao, M. Glidden, D. O'Donnell, C. Tainter, M. Boustani, and A. Alexander, "A novel emergency medical services protocol to improve treatment time for large vessel occlusion strokes," *Plos one,* vol. 17, no. 2, p. e0264539, 2022.

[19] S. Walter, H. Zhao, D. Easton, C. Bil, J. Sauer, Y. Liu, M. Lesmeister, I. Q. Grunwald, G. A. Donnan, and S. M. Davis, "Air-Mobile Stroke Unit for access to stroke treatment in rural regions," *International journal of stroke,* vol. 13, no. 6, pp. 568-575, 2018.

[20] A. Shuaib and T. Jeerakathil, "The mobile stroke unit and management of acute stroke in rural settings," *Cmaj,* vol. 190, no. 28, pp. E855-E858, 2018.

[21] E. Aronsson and A. Hoffmen, *A Dynamic Exploration of Optimal Meeting Points Between Mobile Stroke Units and Ambulances in Prehospital Stroke Care*, Bachelor thesis, Malmö University, 2024.

[22] M. P. Kate, T. Jeerakathil, B. H. Buck, K. Khan, A. Z. Nomani, A. Butt, S. Thirunavukkarasu, T. Nowacki, H. Kalashyan, and M. I. Lloret-Villas, "Pre-hospital triage of suspected acute stroke patients in a mobile stroke unit in the rural Alberta," *Scientific Reports,* vol. 11, no. 1, p. 4988, 2021.

[23] J. Al Fatah, A. a. Alshaban, J. Holmgren, and J. Petersson, "An agent-based simulation model for assessment of prehospital triage policies concerning destination of stroke patients," *Procedia Computer Science,* vol. 141, pp. 405-412, 2018.

[24] V. Bélanger, A. Ruiz, and P. Soriano, "Recent optimization models and trends in location, relocation, and dispatching of emergency medical vehicles," *European Journal of Operational Research,* vol. 272, no. 1, pp. 1-23, 2019.

[25] R. Aringhieri, M. E. Bruni, S. Khodaparasti, and J. T. van Essen, "Emergency medical services and beyond: Addressing new challenges through a wide literature review," *Computers & Operations Research,* vol. 78, pp. 349-368, 2017.

[26] O. Dahllöf, F. Hofwimmer, J. Holmgren, and J. Petersson, "Optimal placement of Mobile Stroke Units considering the perspectives of equality and efficiency," *Procedia Computer Science,* vol. 141, pp. 311-318, 2018.

[27] J. Lundgren, M. Rönnqvist, and P. Värbrand, *Optimization*. Studentlitteratur, 2010.

[28] R. Buter, A. Nazarian, H. Koffijberg, E. W. Hans, R. Stieglis, R. W. Koster, and D. Demirtas, "Strategic placement of volunteer responder system defibrillators," *Health care management science,* pp. 1-22, 2024.

[29] H. Leknes, E. S. Aartun, H. Andersson, M. Christiansen, and T. A. Granberg, "Strategic ambulance location for heterogeneous regions," *European Journal of Operational Research,* vol. 260, no. 1, pp. 122-133, 2017.

[30] M. Lujak, H. Billhardt, and S. Ossowski, "Distributed coordination of emergency medical service for angioplasty patients," *Annals of Mathematics and Artificial Intelligence,* vol. 78, pp. 73-100, 2016.

[31] T. G. Phan, R. Beare, V. Srikanth, and H. Ma, "Googling location for Mobile Stroke Unit hub in metropolitan Sydney," *Frontiers in neurology,* vol. 10, p. 810, 2019.

[32] J. P. Rhudy Jr, A. W. Alexandrov, J. Rike, T. Bryndziar, A. H. Z. Maleki, V. Swatzell, W. Dusenbury, E. J. Metter, and A. V. Alexandrov, "Geospatial visualization of mobile stroke unit dispatches: a method to optimize service performance," *Interventional neurology,* vol. 7, no. 6, pp. 464-470, 2018.

[33] M. A. Abid, S. Amouzad Mahdiraji, F. Lorig, J. Holmgren, R.-C. Mihailescu, and J. Petersson, "A Genetic Algorithm for Optimizing Mobile Stroke Unit Deployment," *Procedia Computer Science,* vol. 225, pp. 3536-3545, 2023.

[34] A. M. Law, W. D. Kelton, and W. D. Kelton, *Simulation modeling and analysis*. McGraw-Hill New York, 2000.

[35] M. N. Abourraja, L. Marzano, J. Raghothama, A. B. Asl, A. S. Darwich, S. Meijer, S. Lethvall, and N. Falk, "A data-driven discrete event simulation model to improve emergency department logistics," in *2022 Winter Simulation Conference (WSC),* 2022: IEEE, pp. 748-759.

[36] B. Vieira, D. Demirtas, J. B van de Kamer, E. W. Hans, and W. van Harten, "Improving workflow control in radiotherapy using discrete-event simulation," *BMC medical informatics and decision making,* vol. 19, pp. 1-13, 2019.

[37] E. Moustaid, R. Richard, and S. Meijer, "Agent-Based Modeling of a Network of Emergency Departments in Urban Environments," in *2018 International Conference on Computational Science and Computational Intelligence (CSCI),* 2018: IEEE, pp. 697-702.

[38] S. Almagooshi, "Simulation modelling in healthcare: Challenges and trends," *Procedia Manufacturing,* vol. 3, pp. 301-307, 2015.

[39] H. Billhardt, M. Lujak, V. Sánchez-Brunete, A. Fernández, and S. Ossowski, "Dynamic coordination of ambulances for emergency medical assistance services," *Knowledge-Based Systems,* vol. 70, pp. 268-280, 2014.

[40] M. Keshtkaran, J. Hearne, B. Abbasi, and L. Churilov, "Stroke care systems: can simulation modeling catch up with the recent advances in stroke treatment?," in *2015 Winter Simulation Conference (WSC),* 2015: IEEE, pp. 1379-1390.

[41] B. M. Bogle, A. W. Asimos, and W. D. Rosamond, "Regional evaluation of the severity-based stroke triage algorithm for emergency medical services using discrete event simulation," *Stroke,* vol. 48, no. 10, pp. 2827-2835, 2017.

[42] E. Bonabeau, "Agent-based modeling: Methods and techniques for simulating human systems," *Proceedings of the national academy of sciences,* vol. 99, no. suppl_3, pp. 7280-7287, 2002.

[43] J. A. Nasir and Y.-H. Kuo, "A decision support framework for home health care transportation with simultaneous multi-vehicle routing and staff scheduling synchronization," *Decision Support Systems,* vol. 138, p. 113361, 2020.

[44] M. K. Traoré, G. Zacharewicz, R. Duboz, and B. Zeigler, "Modeling and simulation framework for value-based healthcare systems," *Simulation,* vol. 95, no. 6, pp. 481-497, 2019.

[45] J. Knight, S. Wells, R. Marshall, D. Exeter, and R. Jackson, "Developing a synthetic national population to investigate the impact of different cardiovascular disease risk management strategies: A derivation and validation study," *PloS one,* vol. 12, no. 4, p. e0173170, 2017.

[46] A. Alassadi, F. Lorig, and J. Holmgren, "An Agent-based Model for Simulating Travel Patterns of Stroke Patients," in *DIGITAL 2021–Advances on Societal Digital Transformation, 14-18 November 2021, Athens, Greece,* 2021: ThinkMind, pp. 11-16.

[47] J. Aked, H. Delavaran, B. Norrving, and A. Lindgren, "Temporal trends of stroke epidemiology in southern Sweden: a population-based study on stroke incidence and early case-fatality," *Neuroepidemiology,* vol. 50, no. 3-4, pp. 174-182, 2018.

[48] Y. Wang, A. G. Rudd, and C. D. Wolfe, "Age and ethnic disparities in incidence of stroke over time: the South London Stroke Register," *Stroke,* vol. 44, no. 12, pp. 3298-3304, 2013.

[49] J. B. Valentin, N. H. Hansen, A. B. Behrndtz, U. Væggemose, and M. F. Gude, "Effect of urgency level on prehospital emergency transport times: a natural experiment," *Internal and emergency medicine,* vol. 19, no. 2, pp. 445-453, 2024.

[50] D. J. O'Brien, T. G. Price, and P. Adams, "The effectiveness of lights and siren use during ambulance transport by paramedics," *Prehospital Emergency Care,* vol. 3, no. 2, pp. 127-130, 1999.

[51] M. Juninger and N. Narvell, *On the use of routing engines for dynamic travel time calculation within emergency vehicle transport simulation*, Bachelor thesis, Malmö University, 2023.

[52] T.-Y. Hu, C.-C. Tong, T.-Y. Liao, and W.-M. Ho, "Simulation-assignment-based travel time prediction model for traffic corridors," *IEEE Transactions on Intelligent Transportation Systems,* vol. 13, no. 3, pp. 1277-1286, 2012.

[53] M. A. Abid, F. Lorig, J. Holmgren, and J. Petersson, "Ambulance travel time estimation using spatiotemporal data," *Procedia Computer Science,* vol. 238, pp. 265-272, 2024.

[54] N. Torres, L. Trujillo, Y. Maldonado, and C. Vera, "Correction of the travel time estimation for ambulances of the red cross Tijuana using machine learning," *Computers in Biology and Medicine,* vol. 137, p. 104798, 2021.

[55] Ľ. Buzna and P. Czimmermann, "On the Modelling of Emergency Ambulance Trips: The Case of the Žilina Region in Slovakia," *Mathematics,* vol. 9, no. 17, p. 2165, 2021.

[56] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design science in information systems research," *MIS quarterly,* pp. 75-105, 2004.

[57] S. T. March and G. F. Smith, "Design and natural science research on information technology," *Decision support systems,* vol. 15, no. 4, pp. 251-266, 1995.

[58] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," *Journal of management information systems,* pp. 45-77, 2007.

[59] Statistics Sweden, "demographic data 2018," [Online]. Available: https://www.scb.se/. [Accessed: Sept. 6, 2018].

[60] Sweden's Southern Regional Health Care Committee, "stroke data 2018," [Online]. Available: https://sodrasjukvardsregionen.se. [Accessed: Sept. 6, 2018].

[61] SOS Alarm, "SOS alarm data for 2020-2021," Available: https://www.sosalarm.se/. [Accessed: Oct. 1, 2023].

[62] M. P. Lin, N. Sanossian, and D. S. Liebeskind, "Imaging of prehospital stroke therapeutics," *Expert review of cardiovascular therapy,* vol. 13, no. 9, pp. 1001-1015, 2015.

[63] S. John, S. Stock, R. Cerejo, K. Uchino, S. Winners, A. Russman, T. Masaryk, P. Rasmussen, and M. S. Hussain, "Brain imaging using mobile CT: current status and future prospects," *Journal of Neuroimaging,* vol. 26, no. 1, pp. 5-15, 2016.

[64] N. E. Vogel, P. Wester, T. A. Granberg, and L.-Å. Levin, "Optimized density and locations of stroke centers for improved cost effectiveness of mechanical thrombectomy in patients with acute ischemic stroke," *Journal of NeuroInterventional Surgery,* vol. 16, no. 2, pp. 156-162, 2024.

[65] N. E. Vogel, P. Wester, T. A. Granberg, and L. Å. Levin, "Cost-Effectiveness of Prehospital Ambulance Helicopter Transportation of Patients With Presumed Stroke in the Era of Mechanical Thrombectomy," *Stroke: Vascular and Interventional Neurology,* vol. 4, no. 5, p. e001343, 2024.

PART II:
PAPERS

PAPER I - MOBILE STROKE UNITS FOR ACUTE STROKE CARE IN THE SOUTH OF SWEDEN

Saeid Amouzad Mahdiraji, Oliver Dahllöf, Felix Hofwimmer, Johan Holmgren, Radu-Casian Mihailescu, and Jesper Petersson

ABSTRACT

A Mobile stroke unit (MSU) is a type of ambulance deployed to promote the rapid delivery of stroke care. We present a computational study using a time to treatment estimation model to analyze the potential benefits of using MSUs in Sweden's Southern Health Care Region (SHR). In particular, we developed two scenarios (MSU1 and MSU2) each including three MSUs, which we compared with a baseline scenario containing only regular ambulances. For each MSU scenario, we assessed how much the expected time to treatment is estimated to decrease for the whole region and each subregion of SHR, and how the population is expected to benefit from the deployment of MSUs. For example, the average time to treatment in SHR was decreased with 20,4 and 15,6 minutes, respectively, in the two MSU scenarios. Moreover, our computational results show that the locations of the MSUs significantly influence what benefits can be expected. While MSU1 is expected to improve the situation for a higher share of the population, MSU2 is expected to have a higher impact on the patients who currently have the longest time to treatment.

PUBLISHED IN

Cogent Engineering, vol. 8, 2021, doi:
https://doi.org/10.1080/23311916.2021.1874084.

1. Introduction

Stroke is a medical condition caused either by a clot (ischemic stroke) or a bleeding (hemorrhagic stroke), leading to reduced blood flow in the vessels inside the brain. The lack of blood flow causes severe damage to the brain cells, and without immediate treatment, the patient has a very low chance to make a satisfactory recovery. According to the World Stroke Organization [1], one out of six persons worldwide is expected to suffer a stroke during their lifetime. Each year, 15 million individuals suffer a stroke, and 5.8 million persons die. According to the Swedish Stroke Register [2], only in Sweden, over 21.000 persons suffered a stroke each year, and in Sweden's Southern health care region (SHR), which is the focus area of the current study, the number is approximately 3900. Stroke is also a leading cause of long-term disability, leaving individuals and families paralyzed and unable to work and provide for each other.

It is generally accepted that the time to treatment is the most important factor for the ability to rehabilitate the stroke patients' mental and physical abilities. Therefore, the term "golden hour" is often used to emphasize the importance of providing early treatment of stroke patients; the patients who are given treatment within an hour have much better chances to recover than those patients where the treatment is initiated later [3]. However, providing fast treatment of stroke patients is far from trivial, which to a large extent is due to logistical challenges and difficulties of determining the correct stroke diagnosis. It must be underlined that different types of stroke need different types of treatment. In particular, ischemic stroke patients should be given intravenous treatment (thrombolysis), which breaks down the clot, whereas this type of treatment under no circumstances should be given to hemorrhagic stroke patients.

In order to identify the type of stroke, a computed tomography (CT) scan needs to be performed; however, as CT scanners are typically only available at the hospitals, it is not possible to initiate treatment until after the patient has been diagnosed at the nearest acute hospital. For stroke patients located far away from an acute hospital, this means that it is typically not possible to provide treatment within the "golden hour".

A mobile stroke unit (MSU) is an ambulance equipped with CT scanner for diagnosis of stroke patients and with personnel who are trained to perform some types of stroke treatment, in particular thrombolysis, which is the standard treatment for ischemic stroke. The use of MSUs has potential benefits also for patients with hemorrhagic stroke by enabling to start blood pressure reducing treatment already in the ambulance. As diagnosis can be provided already in the ambulance, MSUs allow cutting the time for diagnosis, and in most cases the time to treatment, both for ischemic and hemorrhagic stroke patients, corresponding to at least the time needed for transporting the patient to the hospital. MSUs are being used in some cities around the world, e.g., Berlin [3, 4] and Cleveland [5]. Compared to the situation where only regular ambulances are employed for stroke transport logistic, the use of MSUs has led to substantial reduction in the time to treatment [3-5]. Furthermore, the model-based analyses of Reimer et al. [6] and Walter et al. [7] indicate the cost-effectiveness for society when introducing MSUs.

The purpose of our work is to contribute to reduced time to treatment for stroke patients, which is expected to lead to significantly reduced consequences of stroke, both in terms of human suffering and social costs. We extend the work of Dahllöf et al. [8], who study the potential benefits of using MSUs within the Skåne county of Sweden, which is a subregion of SHR. The contribution of the current study is twofold: i) a model for estimating the expected time to treatment for different parts of a geographic region and ii) the results of a case study where we applied our model to SHR. In particular, we compared the current situation in SHR, using only regular ambulances, with two scenarios including three MSUs each. SHR is a large region including both urban and rural areas; and we used a fine-grained division of SHR into subregions in order to take into account the varying density of the population and of the expected time to treatment over different parts of SHR. The results of our study indicate that using only a few MSUs, it is possible to significantly reduce the time to treatment for a large amount of the inhabitants of SHR. The obtained results also have the potential to serve as decision support for policymakers who consider deploying MSUs in order to reduce the time to treatment within their geographical region.

The remainder of the current article is organized as follows. In Section 2, we give an account to the related work. In Section 3, we describe our average time to treatment estimation model used in our scenario study, which is presented in Section 4. Our computational results are presented in Section 5. We provide a discussion on our results in Section 6, and the study is concluded in Section 7.

2. Related Work

The literature contains a large number of studies on improving the health care transport logistics using simulation and optimization modeling, focusing on optimal location, relocation, and dispatching of emergency medical vehicles over a geographical region. In particular, some studies use simulation modeling in order to explicitly assess prehospital stroke policies. Holodinsky et al. [9] employ conditional probability modeling to estimate the effect of treatment times on prehospital transport decision-making for ischemic stroke patients. They compare two different treatment policies, i.e., *drip and ship* and *direct to endovascular thrombectomy*. Schlemm et al. [10] study the consequences of prehospital stroke strategies in order to predict which type of treatment should be given to stroke patients. Al Fatah et al. [11] use agent-based simulation to evaluate two stroke transport policies, i.e., *nearest hospital* and *nearest hospital towards the stroke center*, concerning where to transport potential stroke patients for diagnosis. Sarraj et al. [12] introduce two optimization models, which they deploy in four states of the USA, in order to assess the same policies that Al Fatah et al. study.

There are some recent studies on the optimal placement of an MSU in a geographical region, where the impact of placing an MSU for inhabitants of urban [13, 14] or rural areas [15] are investigated. Rhudy et al. [14] optimize emergency service delivery for stroke patients in the city of Memphis, USA, using geospatial analysis of the distribution of stroke cases. Phan et al. [13] employ the ggmap interface with the Google Maps API to identify the optimal placement of an MSU in Sydney, Australia. Dahllöf et al. [8] use expected value optimization in order to identify the optimal location of an MSU in the Skåne county of Sweden. They propose that MSU locations in the region should be determined considering two distinct perspectives, efficiency and equity. Efficiency aims to place an MSU where a higher proportion of the population is expected to receive a shorter time to treatment, whereas equity emphasizes on reducing the time to treatment for the patients who live farthest from an acute hospital.

While earlier studies mainly have studied the advantages of deploying an MSU in densely populated areas, we applied our model in a larger area, which consists of both urban and rural areas.

3. Average Time to Treatment Model

We here present our model for estimating the average time to treatment for stroke patients in a geographical region. We divide the geographical region under consideration into a non-overlapping set of subregions, denoted by R. The union $\bigcup_{r \in R} r$ of all subregions equals the whole region under consideration. We assumed that all stroke patients residing in subregion $r \in R$ are represented by the centroid $c(r)$ of r. This means that all transport to and from a particular subregion $r \in R$ is made to and from its centroid $c(r)$.

We let L^{AMB} denote the set of regular ambulance locations, L^{MSU} the set of MSU locations, and H the set of acute hospital locations in the region under consideration. For a regular ambulance located at ambulance site $l \in L^{AMB}$, $t_l^{AMB\ RESP}$ denotes the expected time from the emergency call until an ambulance starts driving towards the patient and $t_{lr}^{AMB\ LR}$ is the expected time to drive from location l to the centroid $c(r)$ of subregion r. We let $t^{AMB\ LAY}$ denote the layover time for a regular ambulance, which is the expected time from the arrival of the ambulance to the patient location until it departs. We let $t_{rh}^{AMB\ RH}$ be the expected time to drive from $c(r)$ to acute hospital $h \in H$ and t_h^{DIAG} the expected time for diagnosis at acute hospital $h \in H$. We assume that the treatment can be initiated immediately after diagnosis, at least for the ischemic stroke patients that are eligible for thrombolysis.

For an MSU located at ambulance location $l \in L^{MSU}$, we let $t_l^{MSU\ RESP}$ denote the expected time from the emergency call until an MSU starts driving towards the patient, $t_{lr}^{MSU\ LR}$ the expected time for an MSU to drive from location l to the centroid $c(r)$ of subregion r, $t^{MSU\ LAY}$ the expected layover time, and $t^{MSU\ DIAG}$ the expected time for diagnosis of the patient inside the MSU. The layover time for an MSU is the expected time between the arrival of the MSU to the patient location until diagnosis is initiated.

Considering only the regular ambulances located in L^{AMB}, the expected time to treatment for a patient located in subregion $r \in R$ is calculated as:

$$t_r^{AMB\ TT} = \min_{l \in L, h \in H} \{t_l^{AMB\ RESP} + t_{lr}^{AMB\ LR} + t^{AMB\ LAY} + t_{rh}^{AMB\ RH} + t_h^{DIAG}\}, \qquad (1)$$

where the ambulance is assumed to drive from the closest ambulance location and to the closest acute hospital.

The expected time to treatment using only the MSUs located in L^{MSU}, for a patient located in subregion $r \in R$, is calculated as:

$$t_r^{MSU\ TT} = \min_{l \in L, h \in H} \{t_l^{MSU\ RESP} + t_{lr}^{MSU\ LR} + t^{MSU\ LAY} + t^{MSU\ DIAG}\}. \qquad (2)$$

The expected time to treatment, where both regular ambulances and MSUs are available, for a patient located in subregion $r \in R$, is calculated as:

$$t_r^{TT} = \min \{t_r^{AMB\ TT}, t_r^{MSU\ TT}\}. \qquad (3)$$

It should be emphasized that Equation (3) ensures that the average time to treatment model assumes that either the nearest MSU or the nearest regular ambulance is chosen for each of the subregions $r \in R$, depending on which of the options are expected to lead to the shortest time for diagnosis. For example, if the expected time to treatment using the nearest MSU for a patient located in subregion $r \in R$ is lower than the expected time to treatment using the nearest ambulance, the model assumes that the MSU will be dispatched to the patient location. In particular, this means that the model assumes that the choice of ambulance service is made without taking into consideration a limited dispatch radius for the regular ambulance or for the MSU.

It should be also emphasized that our model aims to minimize the expected time for diagnosis, where we assumed that treatment, that is, thrombolysis for ischemic stroke patients or blood pressure reducing treatment for hemorrhagic stroke patients, can be provided immediately after diagnosis. Please note that our estimations of the time to treatment is based on estimations of the time for diagnosis of ischemic stroke patients. In summary, the model assumes that, when a patient is picked up, the regular ambulance departs from the patient location to the closest acute hospital for diagnosis. For the MSU dispatches, the model assumes that treatment can be initiated, inside the MSU, immediately after diagnosis, before transporting the patient to an acute hospital.

Let Q_r denote the share of the stroke cases in the considered region that is expected to occur in subregion $r \in R$, where $\sum_{r \in R} Q_r = 1$. The weighted average time to treatment for the whole region is:

$$t^{TT} = \sum_{r \in R} t_r^{TT} \cdot Q_r. \tag{4}$$

The average, non-weighted, time to treatment for the whole region is:

$$t^{TT\,AVG} = \frac{1}{|R|} \sum_{r \in R} t_r^{TT}. \tag{5}$$

It is also possible to calculate the weighted average time to treatment for any set of subregions $R' \subset R$, which can refer to a municipality or any other region of interest. The weighted average time to treatment for any subset $R' \subset R$ is:

$$t_{R'}^{TT} = \sum_{r \in R'} t_r^{TT} \cdot \left(\frac{Q_r}{\sum_{r \in R'} Q_r} \right). \tag{6}$$

The average (non-weighted) time to treatment taken over all subregions in $R' \subset R$ is:

$$t_{R'}^{TT\ AVG} = \frac{1}{|R'|} \sum_{r \in R'} t_r^{TT}. \tag{7}$$

4. Scenario Description

In this section, we describe our application of the presented average time to treatment model for SHR. Furthermore, we describe the considered assumptions and the data processing conducted within our scenario study.

4.1. Sweden's Southern Health Care Region

As mentioned above, we considered Sweden's Southern Health Care Region (SHR), which consists of 49 municipalities located in the four counties of Blekinge, Halland, Kronoberg, and Skåne (see Table A. 1 for more information), and which covers an area of 27.900 square kilometers. The population of SHR was 1.687.000 in 2018, where Malmö (Skåne County), Lund (Skåne County), and Halmstad (Halland County) are the largest populated municipalities with a population of 322.000, 122.000, and 98.800, respectively [16]. There are 13 acute hospitals equipped with a CT scanner in SHR, and ambulances are located at 39 ambulance sites. Due to the geographical limitation of our study, we chose to not consider islands without fixed connection to the mainland. See Figure 1 for an overview of SHR, where the locations of the ambulance sites and acute hospitals are represented by green triangles and purple circles, respectively. In Figure 1, we also illustrate the population distribution over the region, where the darker the square color is, the higher the population density is.

In 2018, about 3900 stroke cases was registered in southern Sweden, where approximately 20% of the cases were hemorrhagic cases and approximately 80% were ischemic stroke cases. About 20% of the ischemic stroke cases received recanalization therapy, including thrombolysis, both thrombolysis and thrombectomy, or only thrombectomy [2]. One of the main reasons why about 80% of the ischemic stroke cases do not get treatment is that they arrive too late to the acute hospital. By reducing the time to treatment, recanalization therapy could potentially be given to a much larger share of the ischemic stroke patients.

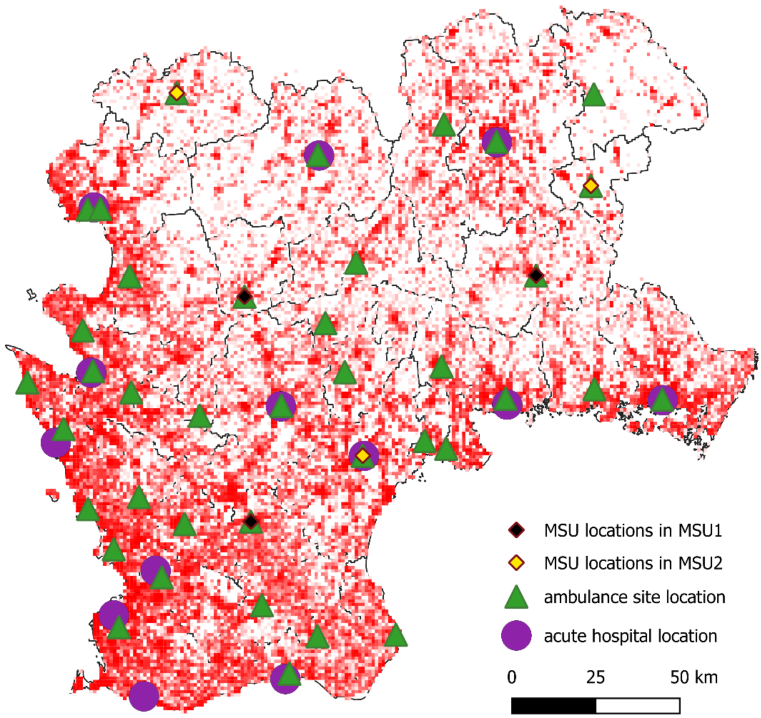


Figure 1. Overview of SHR, where the green triangles and purple circles rep-resent the locations of ambulance sites and acute hospitals, respectively. The MSUs in the MSU1 and MSU2 scenarios are presented by black and yellow diamonds, respectively. The color of each square displays its population density. The lighter squares have lower density.

4.2. Scenarios

In our scenario study, we analyzed the possible implications of placing MSUs at different locations in SHR. In particular, we considered three scenarios: A baseline scenario containing only regular ambulances and two extended scenarios containing three MSUs each. The baseline scenario corresponds to the current situation, where regular ambulances are located in all of the 39 ambulance sites shown in Figure 1. The baseline scenario enabled us to identify areas of SHR that are problematic considering the expected time to treatment for stroke patients.

We assumed that all of the 39 ambulance sites located in SHR are candidate locations for MSUs. By visually analyzing the results for the baseline scenario, we created two extended scenarios, which in addition to the regular ambulances, include three MSUs each:

- MSU1: MSUs located in Hörby, Markaryd, and Tingsryd
- MSU2: MSUs located in Lessebo, Hyltebruk, and Kristianstad

In Figure 1, we illustrate the MSU locations in MSU1 and MSU2 scenarios by black and yellow diamonds, respectively. When we created the two MSU scenarios, we made a tradeoff between the number of MSUs and the coverage of SHR. In particular, we visually analyzed the map in Figure 3a (see Section 5), where we chose to place the MSUs so that they appear to cover as much of the problematic areas, that is, areas with long expected time to treatment, as possible, and not being too close to any acute hospital. The reason for considering scenarios with no more than three MSUs was that this appears to provide decent coverage of SHR; we further took into consideration that the operational costs for an MSU are high.

In our scenario, we assumed that all dispatches of MSUs or regular ambulances are for actual stroke cases. In addition, due to limitations of the data that was available for our study, we used expert knowledge in order to make the following assumptions:

- It takes three minutes from an emergency call until an ambulance starts driving towards the patient site ($t_l^{AMB\ RESP} = t_l^{MSU\ RESP} = 0{,}05$h).
- The layover time for both regular ambulances and MSUs is 15 minutes regardless of where the patient is located ($t^{AMB\ LAY} = t^{MSU\ LAY} = 0{,}25$h).
- The time for diagnosis of a patient inside an MSU is 15 minutes ($t^{MSU\ DIAG} = 0{,}25$h).
- The time for diagnosis in hospital is 35 minutes for all of the considered acute hospitals ($t_h^{DIAG} = 0{,}583$h). We based the time for diagnosis in hospital on the door-to-needle time for ischemic stroke patients. To the best of our knowledge, the time to in-hospital thrombolysis is around 30 minutes in SHR. In addition, in a conducted study in a Norwegian stroke centre, the door-to-needle time to thrombolysis treatment was reduced from 27 to 13 minutes [17]. Therefore, we argue that 35 minutes is a quite reasonable time for diagnosis in hospital.
- All of the considered ambulance sites and hospitals are open and provide service on 24/7 basis.
- There is always an ambulance or MSU available when it is needed.
- The MSUs in the two MSU scenarios are able to provide service in the whole region, that is, they are not limited in how far they are able to drive.
- All stroke cases occur at the patients' homes.

4.3. Data Processing

We collected the demographic data used in the study from Statistics Sweden [16] and the age-based stroke data was provided by the Sweden's Southern Regional Health Care Committee [18]. In addition, we made a geographic division of SHR into a disjoint set of 1×1 km squares (i.e., R), as this enabled us to study in detail the variation of demography and the expected time to treatment within SHR.

In order to generate the necessary input for the calculations of the weighted average time to treatment for SHR and for the municipalities in SHR (see Section 3), we had to estimate the expected share of stroke cases Q_r for each of the 1×1 km squares $r \in R$ representing SHR. We used the age-based stroke data, together with the demographic data for this aim. The considered stroke data specifies the number of reported stroke cases for 21 age groups over the whole SHR, where we let A= {[0,4), [4,8), …, [95,99), [100, ∞)} denote our disjoint set of age groups. In addition, the demographic data contains the number of inhabitants for each of the age groups and each of the considered 1×1 km squares.

We let pop_{ar} denote the number of inhabitants of age group $a \in A$ in square $r \in R$. The total number of inhabitants for age group a in the whole SHR is given by $pop_a = \sum_{r \in R} pop_{ar}$. For the whole SHR, S_a denotes the number of reported stroke incidents during some specific time period, in our study 2018, for age group $a \in A$. For each of the age groups $a \in A$, we calculated the likelihood that a person will get a stroke during the considered time period as:

$$P(a) = S_a / pop_a. \tag{8}$$

For each of the considered age groups, we present in Figure 2 the calculated likelihoods, i.e., the $P(a)$:s, of an individual getting a stroke during one year.

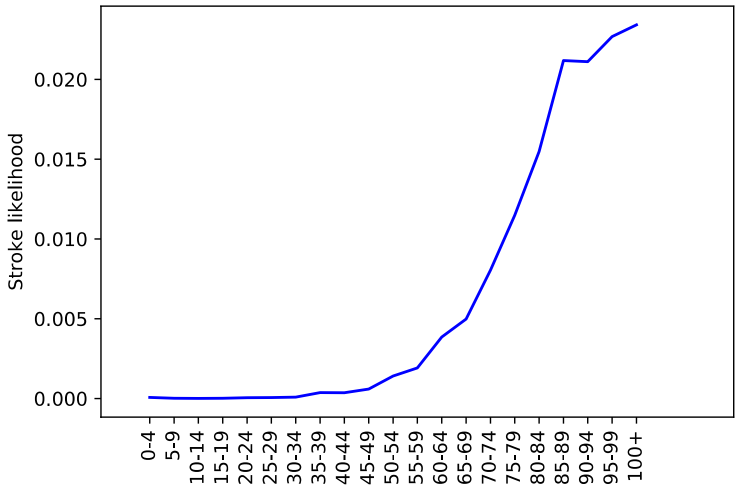


Figure 2. Stroke likelihoods for the considered age groups.

Using the calculated likelihoods, i.e., the $P(a)$:s, and the pop_{ar}:s, we calculated the expected number of stroke incidents S_r for each of the subregions $r \in R$ for our considered time period as:

$$S_r = \sum_{a \in A} P(a) \cdot pop_{ar}. \tag{9}$$

The expected number of stroke cases in the whole region (i.e., SHR) is:

$$S = \sum_{r \in R} S_r. \tag{10}$$

The share of stroke cases that are expected to occur in subregion $r \in R$ is calculated as:

$$Q_r = S_r / S. \tag{11}$$

Please note that Q_r can be also interpreted as the likelihood that the next stroke in the region will occur in subregion r.

We estimated the driving times of both regular ambulances and MSUs using the driving time generation functionality of Open Street Map[1] (OSM), which we

[1] see openstreetmap.org

accessed using the Openrouteservice toolbox in QGIS. We made our calculations using MATLAB 2019B.

For each of the considered ambulance locations $l \in L$, each of the subregions $r \in R$, and each of the acute hospital $h \in H$, we generated the following driving times:

- $t_{lr}^{AMB\_LR}$: The estimated time for a regular ambulance to drive from ambulance location $l \in L$ to the centroid $c(r)$ of $r \in R$.
- $t_{lr}^{MSU\_LR}$: The estimated time for an MSU to drive from ambulance location $l \in L$ to the centroid $c(r)$ of $r \in R$.
- $t_{rh}^{AMB\_RH}$: The estimated time for a regular ambulance to drive from the centroid $c(r)$ of $r \in R$ to the acute hospital $h \in H$.

It should be noted that OSM generates driving times for cars. As we did not have access to ambulance driving times, we instead estimated the ambulance driving times between different parts of SHR using the generated driving times for cars. In fact, we made a conservative assessment of ambulance driving times. According to ambulance personnel in Southern Sweden (personal report), they are instructed to drive with a safety first principle and will thus keep close to official speed limits. Therefore, we assumed that regular ambulances drive 5% faster than a passenger car, and MSUs drive at the same speed as a car. When we made this assumption, we took into consideration the fact that an MSU equipped with a CT scanner is generally heavier and larger than a regular ambulance; hence, it potentially travels slower.

5. Computational Results

In Table 1, we present the weighted and non-weighted average time to treatment for the whole SHR, which we calculated using Equation (4) and Equation (5), respectively, for each of our three scenarios. The average, non-weighted, time to treatment for the whole SHR is the mean of expected times to treatment taken over all of the subregions $r \in R$. The expected time to treatment refers to the elapsed time from receiving a call of a stroke case until the treatment is initiated. The weighted average time to treatment for the whole region is the sum of the expected times to treatment taken over all subregions $r \in R$, where the expected time to treatment for subregion r is weighted by Q_r, that is, the likelihood that a stroke occurs in subregion r. Please see Section 3 for more details. The shortest and longest expected time to treatment for each of the scenarios are presented in Table 2.

Table 1. Weighted and non-weighted average time to treatment for the whole SHR and all of the considered scenarios. Please note that for both non-weighted and weighted average time to treatment, the expected time to treatment is the time from receiving a call until the treatment is initiated. Please see Section 3 for more details.

Scenario	Non-weighted average time to treatment	Weighted average to treatment
Baseline	1,62h	1,33h
MSU1	1,28h	1,18h
MSU2	1,36h	1,22h

Table 2. The shortest and longest expected time to treatment for any square for each of the three scenarios.

Scenario	Shortest expected time to treatment	Longest expected time to treatment
Baseline	0,91h (Sjöbo)	2,82h (Hylte)
MSU1	0,57h (Hörby)	2,29h (Uppvidinge)
MSU2	0,56h (Lessebo)	2,27h (Vellinge)

In Figure 3, we present for each of our three scenarios, the expected time to treatment for each of the 1×1 km squares of SHR. Please note that the black dots in sub-figures (b) and (c) of Figure 3 show the locations of the MSUs in the MSU1 and MSU2 scenarios, respectively (see also Figure 1). It should be noted that the lighter the color is in the figures, the shorter the expected time to treatment is for the corresponding square. The results in Figure 3(b) and 3(c) can be also interpreted as how each MSU in each of the two MSU scenarios is expected to impact the expected time to treatment for the different parts of SHR.

a)

b)

c)

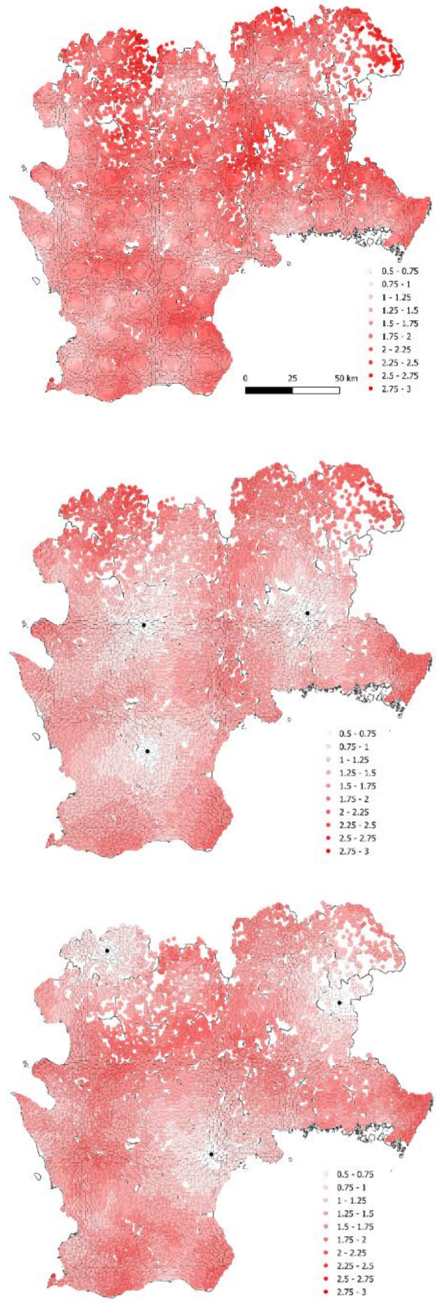


Figure 3. Expected time to treatment for each of the 1×1 km squares covering SHR considering a) only regular ambulances, b) MSU1, and c) MSU2.

In Figure 4, we illustrate, for each of our three scenarios, how the expected time to treatment is expected to vary for the 1×1 km squares covering SHR. It should be noted that the more left-skewed the histogram is, the larger part of SHR is expected to get shorter time to treatment. In particular, this can be seen for MSU2 in Figure 4c. In each of the histograms, the y-axis is the number of squares for each of the 3-minute time to treatment intervals shown on the x-axis.

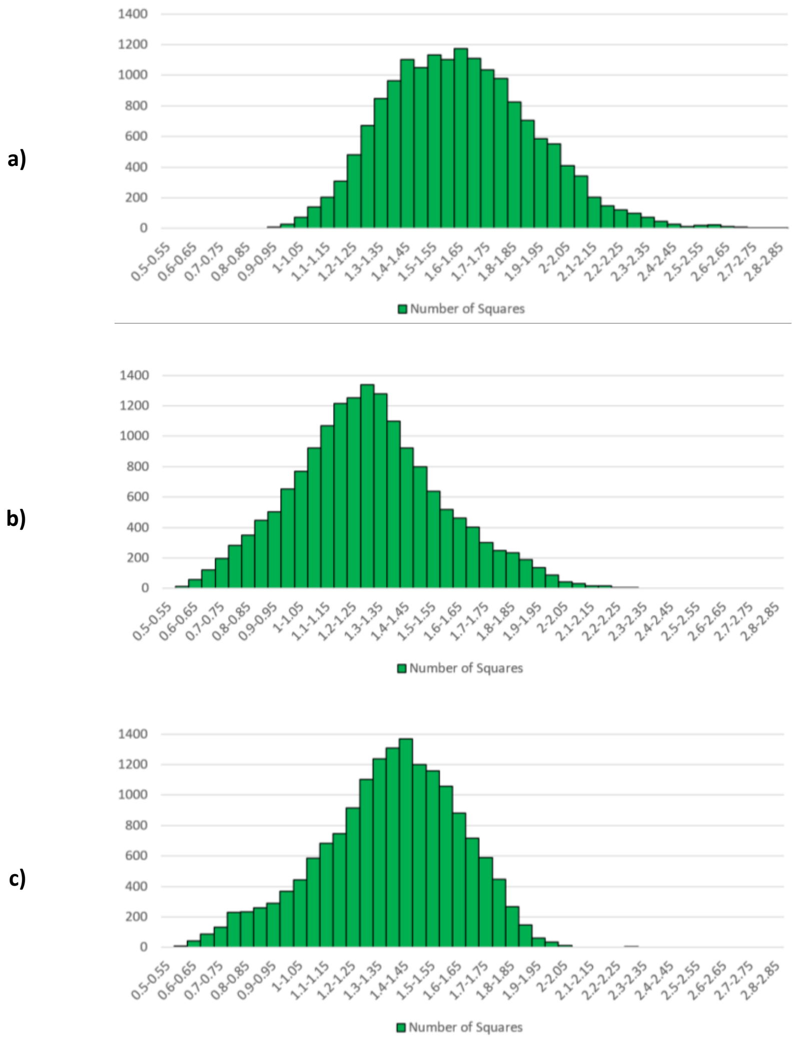


Figure 4. Histogram showing the distribution of the expected time to treatment over the 1×1 km squares of SHR for a) the baseline scenario, b) MSU1, and c) MSU2.

As the histograms in Figure 4 show the number of squares with different expected time to treatment, they do not directly inform about how many of the inhabitants of SHR are expected to get a better situation as a result of placing the MSUs according to our two MSU scenarios. Instead, they inform about the geographic share of the region that is expected to benefit from the use of the MSUs. In Figure

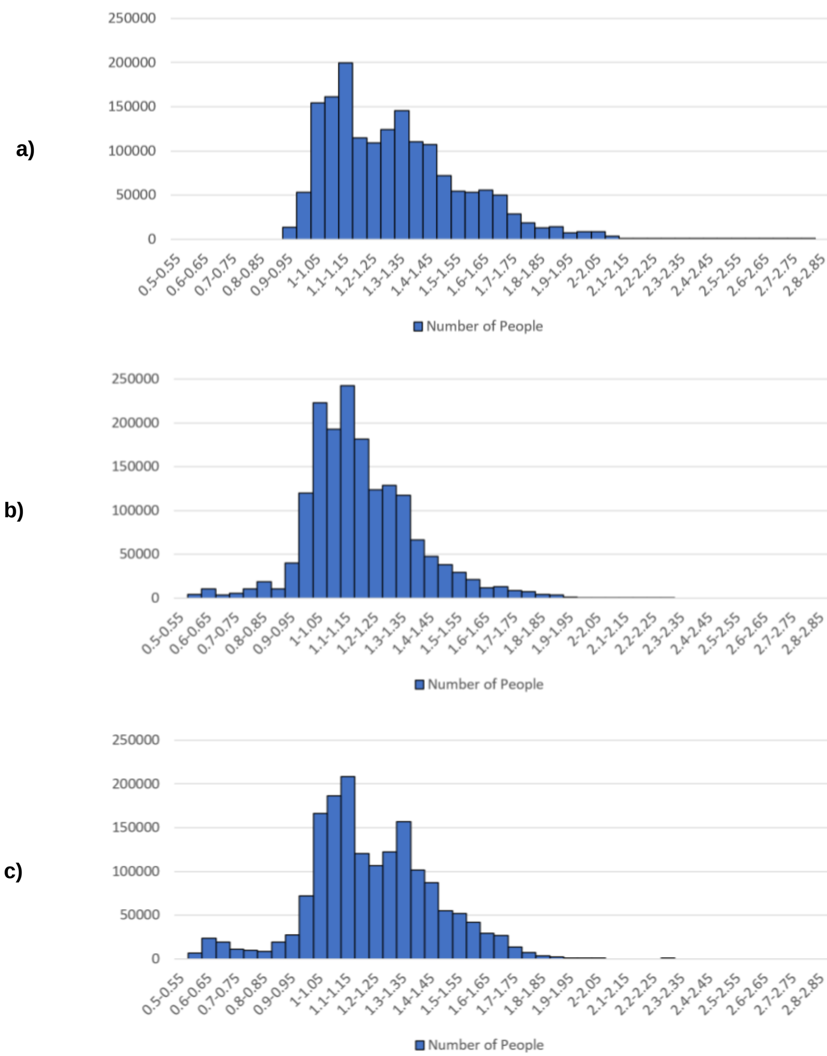


Figure 5. Histogram showing the distribution of the expected time to treatment over the population of SHR for a) the baseline scenario, b) MSU1, and c) MSU2. Each of the bars represents a time interval of 3 minutes (presented in hours).

5, we therefore provide histograms showing how the expected time to treatment is distributed over the population of SHR for our three scenarios. The more left-skewed a histogram is, the more stroke patients are expected to receive treatment earlier. In each of the histograms, the y-axis is the number of inhabitants for each of the 3-minute time to treatment intervals shown on the x-axis.

From the histograms in Figure 5, it is possible to read out the share of inhabitants that are expected to receive treatment within different time thresholds. In order to facilitate the presentation of our results, we provide in Table 3 the share of the total population of SHR (i.e., 1.687.000), that are expected to receive treatment within 60, 75, and 90 minutes for each of the considered scenarios. See also Figure 6, where we present the cumulative distribution function, for each of our three scenarios, of the expected time to treatment for all of the inhabitants of SHR.

Table 3. Share of the inhabitants in SHR that are expected to receive treatment within 60, 75, and 90 minutes for each of the considered scenarios.

Expected time to treatment	**Scenario**		
	Baseline	**MSU1**	**MSU2**
< 60 minutes	3,96%	13,3%	11,7%
< 75 minutes	47,8%	70,4%	58,5%
< 90 minutes	80,9%	94,0%	89,5%

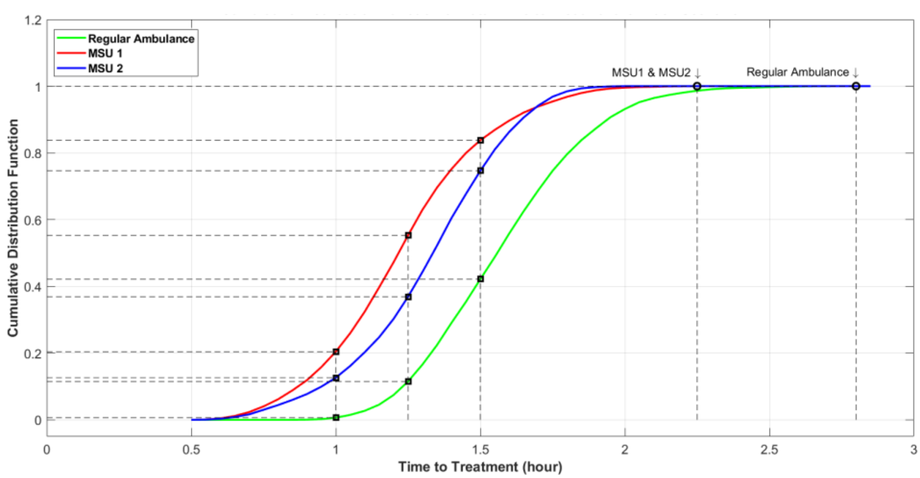


Figure 6. Cumulative distribution function, for each of our three scenarios, of the expected time to treatment for the inhabitants of SHR.

In Figure 7, we show how many of the 1×1 km squares in SHR that are expected to get better service after allocating MSUs according to MSU1 and MSU2. According to the chart, the time to treatment is expected to decrease with up to 100 minutes. In the chart, the x-axis is the time interval (in minutes), and the y-axis is the number of squares.

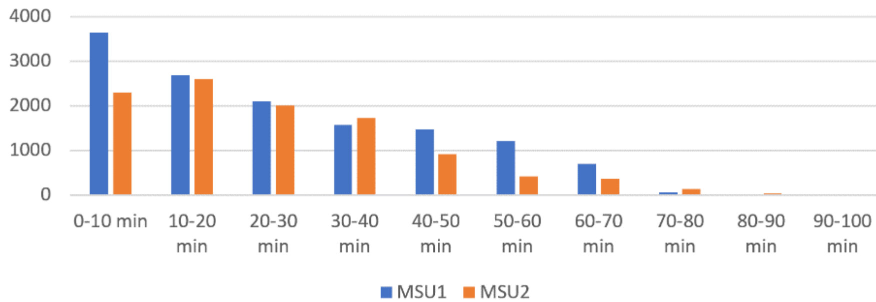


Figure 7. Illustration of how the expected time to treatment is expected to decrease over the 1×1 km squares of SHR according to the MSU1 and MSU2 scenarios.

In Figure 8, we show many of the inhabitants of SHR are expected to receive faster treatment as a result of allocating MSUs according to our two MSU scenarios. In other words, we show how the expected time to treatment is expected to be reduced over the population of SHR for the MSU1 and MSU2 scenarios. In the chart, the x-axis is the time interval (in minutes), and the y-axis is the number of inhabitants.

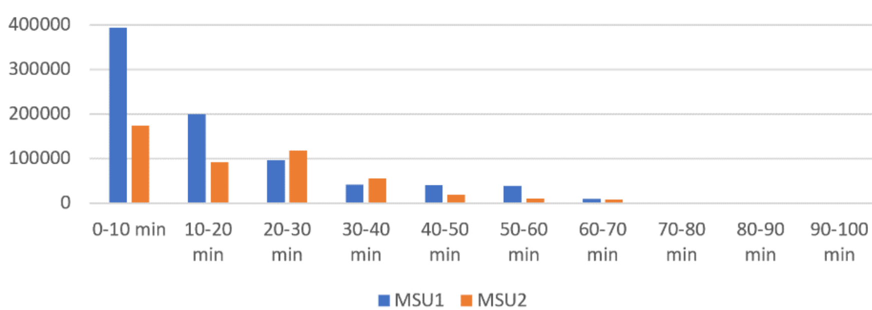


Figure 8. Illustration of how the expected time to treatment is expected to decrease over the population of SHR for the MSU1 and MSU2 scenarios.

Finally, in Figure 9, we show the average time to treatment for each municipality of SHR in the form of a scatter plot. It can be observed that the MSU1 and MSU2 scenarios are able to provide treatment within 60 minutes for patients living in the municipalities Markaryd, Tingsryd, Eslöv, Hörby, Höör, and Örkeljunga; and Hylte, Lessebo, and Kristianstad, respectively. In Table A, we present the weighted and non-weighted average time to treatment for each of the municipalities and for each of the three scenarios. The weighted and non-weighted average times to treatment for each municipality were calculated using Equation (6) and Equation (7), respectively, where R' corresponds to the squares of the considered municipality.

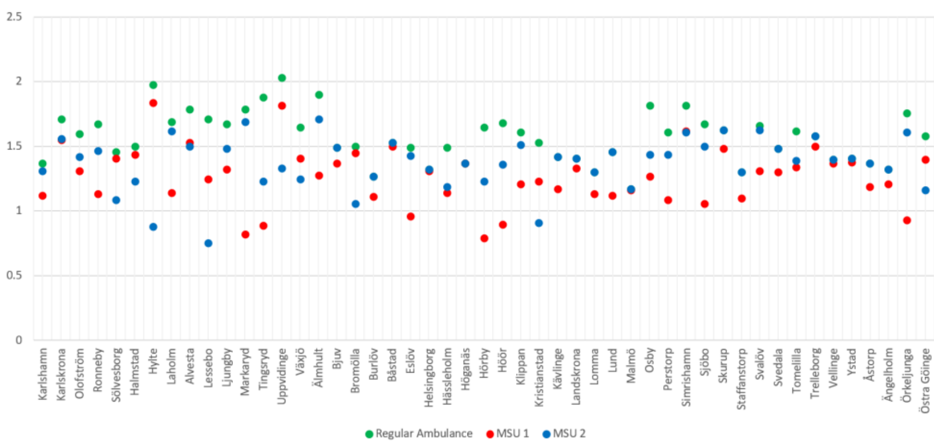


Figure 9. Average time to treatment for each municipality of SHR and for each of the three scenarios.

6.Discussion

The results of our study suggest that using only a few MSUs, the expected time to treatment is expected to significantly decrease for a large share of the inhabitants in SHR. This is emphasized, e.g., by the calculated weighted and non-weighted average times to treatment for our three scenarios. Compared to the baseline scenario, the average time to treatment dropped from 1,62h to 1,28h and 1,36h, respectively for MSU1 and MSU2. The shortest expected time to treatment for any inhabitant in SHR decreased from 0,91h in the baseline scenario to 0,57h in MSU1 and 0,56h in MSU2, and the longest expected time to treatment was reduced from 2,82h to 2,29h in MSU1 and to 2,27h in MSU2 (see Table 2). In addition, the weighted average time to treatment, which was weighted over the population density, was reduced from 1,33h to 1,18h for MSU1 and to 1,22h for MSU2. It should be emphasized here that the reason that the weighted average time to treatment is much lower than the average time to treatment is that the density of people in SHR is higher in those parts that are closer to the ambulance locations and acute hospitals. Visually, this is also shown in Figure 3, where the maps corresponding to MSU1 and MSU2 are much lighter than the map corresponding to the baseline scenario.

By comparing the histograms in Figure 4, it can be further seen that the number of 1×1 km squares with shorter expected time to treatment are significantly higher in the MSU1 and MSU2 scenarios. By comparing the histograms in Figure 5, it can be seen that the number of persons with shorter expected time to treatment is also significantly higher in the MSU1 and MSU2 scenarios. From the figures in these histograms, it is straightforward to read out how large share (or how many) of the population of SHR is expected to receive treatment within a certain time. For example, the share of the population that has an expected time to treatment less than 1h is 3,96%, 13,3%, and 11,7%, respectively for the baseline, MSU1, and MSU2 scenarios. For 1,25 hours (i.e., 75 min), the percentages are 47,8%, 70,4%, and 58,5%, and for 1,5h (i.e., 90 min), the percentages are 80,9%. 94,0%, and 89,5%. Furthermore, for the MSU1 and MSU2 scenarios, 3,35 and 2,96 times as many inhabitants are expected to be treated within 60 minutes compared to the baseline scenario.

According to Figure 7, the inhabitants of 81% and 63% of the 1×1 km squares covering SHR are expected to receive faster treatment according to the MSU1 and MSU2 scenarios, respectively. In addition, the time to treatment of the inhabitants of 63% and 66% of the squares is expected to be reduced up to 30 minutes for MSU1 and MSU2, respectively. Figure 8 shows that the time to treatment is expected to decrease for 82% and 48% of the total population of SHR for the MSU1 and MSU2 scenarios, respectively. In particular, the time to treatment of 84% and 80% of the inhabitants is expected to decrease up to 30 minutes for the MSU1 and MSU2 scenarios, respectively.

It should be noted that the two considered MSU scenarios are two examples of where to locate MSUs. There are other possibilities to place MSUs, but from the results, it can be seen that the different placements of MSUs are expected to lead to benefits for different parts of SHR, which is explicitly shown in Table A, where the weighted and non-weighted average times to treatment are presented for each of the municipalities of SHR. By comparing the results for the two MSU scenarios, we conclude that the MSUs in the MSU1 scenario are expected to improve the situation for a larger share of the population of SHR, whereas the

MSUs in the MSU2 scenario are expected to improve the situation for a larger share of those patients which today has the longest time to treatment. The reason for this is most likely that one of the MSUs in MSU1 is located rather close to the Malmö-Lund region, which is the part of SHR with the highest population density. Hence, it is important to carefully consider the trade-off between efficiency and equity when deciding where to place MSUs.

In the present study, we chose to use *time for diagnosis*. We reasoned that confirmation of a diagnosis of either ischemic or hemorrhagic stroke combining clinical assessment with a CT of the brain is a key decision. In case of an ischemic stroke, treatment with a thrombolytic drug can be started without delay. Furthermore, transportation to a primary or secondary hospital can be planned. If a hemorrhagic stroke is detected, urgent measures like early blood pressure control can be started immediately, and, again, a decision of transportation to a primary or secondary hospital can be made. In trained stroke centers, including those who operate an MSU, the actual additional time from final diagnostic decision to treatment (needle time) is very small and may at most be a few minutes.

7.Conclusions and Future Work

We have presented a computational study focusing on the potential benefits of using mobile stroke units (MSUs) in Sweden's Southern health care region (SHR). In particular, we compared the current situation, where regular ambulances, but no MSUs are located at the 39 ambulance sites within SHR, with two extended scenarios, where we added three MSUs in each.

The results of our study show that the time to treatment is expected to significantly decrease for the two MSU scenarios. In particular, we obtained an improvement of about a factor of 3 when considering the share of the population that is expected to receive treatment within an hour, which can be seen in Table 3. The potential benefit of using MSUs can be exemplified by the municipality of Älmhult (with a population of 17400), where a significant amount of people could be affected by reducing the average time to treatment from 1,9h to 1,28h in the MSU1 scenario.

One of the most important strengths of the study is that we divided SHR into a detailed grid of 1×1 km squares, which enabled us to study in detail how the expected time to treatment is expected to vary within SHR. Another strength we want to highlight is that we used a well-established framework to estimate driving times between the different parts of SHR.

In this study, we focused on the time for diagnosis for stroke patients. However, for ischemic stroke patients, which represent the majority of stroke patients, this is the same as studying the time to treatment for those patients who are eligible for thrombolysis, which in this case is the standard treatment and can be initiated almost immediately after the diagnosis. The ischemic patients with larger clots typically benefit from thrombectomy treatment, which is provided only in Lund, in the southwest of SHR. Since the expected benefits of MSUs are significant concerning the time to thrombolysis, it should be emphasized that the expected benefits for thrombectomy patients are even higher owing to the fact that there would be time-savings corresponding to at least part of the driving time from the patient location to the hospital and from the acute hospital to the thrombectomy center plus the acute hospital layover time. Instead, the MSU can drive directly to the thrombectomy center. This is something that should be further investigated in future work. Future work also includes developing an optimization model to solve the search problem of where it is best to place MSUs and to identify a proper trade-off between the equity and efficiency perspectives. It should be further noted that we did not consider co-dispatching of an MSU with a regular ambulance in the presented model. As part of future work, we plan to apply relevant collaborative policies involving MSUs and regular ambulances of ambulance resources to our model.

Appendix

Table A. Weighted and non-weighted average time (in hours) to treatment for all of the municipalities in SHR and all of the considered scenarios.

County	Municipality	Average time to treatment			Weighted average time to treatment		
		Baseline	**MSU1**	**MSU2**	**Baseline**	**MSU1**	**MSU2**
Blekinge	Karlshamn	1,37	1,12	1,31	1,13	1,08	1,12
	Karlskrona	1,71	1,55	1,56	1,38	1,33	1,34
	Olofström	1,60	1,31	1,42	1,47	1,35	1,33
	Ronneby	1,67	1,13	1,47	1,44	1,18	1,39
	Sölvesborg	1,46	1,41	1,09	1,41	1,38	1,05
Halland	Halmstad	1,50	1,44	1,23	1,18	1,17	1,13
	Hylte	1,98	1,84	0,88	1,79	1,71	0,77
	Laholm	1,69	1,14	1,62	1,49	1,17	1,46

Kronoberg	Alvesta	1,79	1,53	1,50	1,55	1,44	1,39
	Lessebo	1,71	1,25	0,75	1,58	1,22	0,70
	Ljungby	1,67	1,32	1,48	1,27	1,14	1,20
	Markaryd	1,79	0,82	1,69	1,64	0,71	1,62
	Tingsryd	1,88	0,89	1,23	1,78	0,81	1,22
	Uppvidinge	2,03	1,82	1,33	1,92	1,76	1,28
	Växjö	1,65	1,41	1,25	1,28	1,21	1,12
	Älmhult	1,90	1,28	1,71	1,73	1,26	1,61
Skåne	Bjuv	1,49	1,37	1,49	1,43	1,34	1,43
	Bromölla	1,50	1,45	1,06	1,36	1,35	0,97
	Burlöv	1,27	1,11	1,27	1,27	1,12	1,27
	Båstad	1,53	1,50	1,53	1,52	1,47	1,52
	Eslöv	1,49	0,96	1,43	1,38	0,97	1,36
	Helsingborg	1,32	1,31	1,32	1,19	1,19	1,19
	Hässleholm	1,49	1,14	1,19	1,25	1,09	1,07
	Höganäs	1,37	1,37	1,37	1,34	1,34	1,34
	Hörby	1,65	0,79	1,23	1,47	0,68	1,20
	Höör	1,68	0,90	1,36	1,65	0,82	1,33
	Klippan	1,61	1,21	1,51	1,52	1,20	1,47
	Kristianstad	1,53	1,23	0,91	1,30	1,16	0,77
	Kävlinge	1,42	1,17	1,42	1,37	1,16	1,37
	Landskrona	1,41	1,33	1,41	1,36	1,33	1,36
	Lomma	1,30	1,13	1,30	1,36	1,16	1,36
	Lund	1,46	1,12	1,46	1,20	1,03	1,20
	Malmö	1,17	1,16	1,17	1,12	1,11	1,12
	Osby	1,82	1,27	1,44	1,63	1,17	1,36
	Perstorp	1,61	1,09	1,44	1,39	1,17	1,35
	Simrishamn	1,82	1,62	1,61	1,76	1,64	1,62
	Sjöbo	1,67	1,06	1,50	1,54	1,04	1,46
	Skurup	1,63	1,48	1,63	1,59	1,49	1,59
	Staffanstorp	1,30	1,10	1,30	1,29	1,08	1,29
	Svalöv	1,66	1,31	1,63	1,57	1,28	1,55
	Svedala	1,48	1,30	1,48	1,39	1,27	1,39
	Tomelilla	1,62	1,34	1,39	1,45	1,30	1,30
	Trelleborg	1,58	1,50	1,58	1,40	1,38	1,40
	Vellinge	1,40	1,37	1,40	1,48	1,41	1,48
	Ystad	1,41	1,38	1,41	1,22	1,22	1,22
	Åstorp	1,37	1,19	1,37	1,34	1,18	1,34
	Ängelholm	1,32	1,21	1,32	1,09	1,08	1,09
	Örkeljunga	1,76	0,93	1,61	1,68	0,91	1,56
	Östra Göinge	1,58	1,40	1,16	1,48	1,38	1,07

References

[1] World Stroke Organization, "Facts and figures about stroke," [Online]. Available: https://www.world-stroke.org/world-stroke-day-campaign/whystroke-matters/learn-about-stroke. [Accessed: Dec. 20, 2019].

[2] The Swedish Stroke Register, "Stroke registrations," [Online]. Available: https://www.riksstroke.org/sve/forskning-statistikoch-verksamhetsutveckling/statistik/registreringar. [Accessed: Dec. 20, 2019].

[3] M. Ebinger, B. Winter, M. Wendt, J. E. Weber, C. Waldschmidt, M. Rozanski, A. Kunz, P. Koch, P. A. Kellner, and D. Gierhake, "Effect of the use of ambulance-based thrombolysis on time to thrombolysis in acute ischemic stroke: a randomized clinical trial," *Jama*, vol. 311, no. 16, pp. 1622-1631, 2014.

[4] H. Zhao, S. Coote, D. Easton, F. Langenberg, M. Stephenson, K. Smith, S. Bernard, D. A. Cadilhac, J. Kim, and C. F. Bladin, "Melbourne Mobile Stroke Unit and Reperfusion Therapy: Greater Clinical Impact of Thrombectomy Than Thrombolysis," *Stroke*, vol. 51, no. 3, pp. 922-930, 2020.

[5] R. Cerejo, S. John, A. B. Buletko, A. Taqui, A. Itrat, N. Organek, S. M. Cho, L. Sheikhi, K. Uchino, and F. Briggs, "A mobile stroke treatment unit for field triage of patients for intraarterial revascularization therapy," *Journal of Neuroimaging*, vol. 25, no. 6, pp. 940-945, 2015.

[6] P. Reimer, A. Zafar, F. M. Hustey, D. Kralovic, A. N. Russman, K. Uchino, M. S. Hussain, and B. L. Udeh, "Cost-Consequence Analysis of Mobile Stroke Units vs. Standard Prehospital Care and Transport," *Frontiers in Neurology*, vol. 10, p. 1422, 2020.

[7] S. Walter, I. Q. Grunwald, S. A. Helwig, A. Ragoschke-Schumm, M. Kettner, M. Fousse, M. Lesmeister, and K. Fassbender, "Mobile stroke units-cost-effective or just an expensive hype?," *Current atherosclerosis reports*, vol. 20, no. 10, p. 49, 2018.

[8] O. Dahllöf, F. Hofwimmer, J. Holmgren, and J. Petersson, "Optimal placement of Mobile Stroke Units considering the perspectives of equality and efficiency," in *Proceedings of the 8th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH 2018)*, vol. 141, pp. 311-318, 2018.

[9] J. K. Holodinsky, A. B. Patel, J. Thornton, N. Kamal, L. R. Jewett, P. J. Kelly, S. Murphy, R. Collins, T. Walsh, and S. Cronin, "Drip and ship versus direct to endovascular thrombectomy: the impact of treatment times on transport decision-making," *European stroke journal*, vol. 3, no. 2, pp. 126-135, 2018.

[10] L. Schlemm, M. Ebinger, C. H. Nolte, and M. Endres, "Impact of prehospital triage scales to detect large vessel occlusion on resource utilization and time to treatment," *Stroke*, vol. 49, no. 2, pp. 439-446, 2018.

[11] J. Al Fatah, A. a. Alshaban, J. Holmgren, and J. Petersson, "An agent-based simulation model for assessment of prehospital triage policies concerning destination of stroke patients," in *Proceedings of the 8th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH 2018)*, vol. 141, pp. 405-412, 2018.

[12] Sarraj, S. Savitz, D. Pujara, H. Kamal, K. Carroll, F. Shaker, S. Reddy, K. Parsha, L. E. Fournier, and E. M. Jones, "Endovascular Thrombectomy for Acute Ischemic Strokes: Current US Access Paradigms and Optimization Methodology," *Stroke*, vol. 51, no. 4, pp. 1207-1217, 2020.

[13] T. G. Phan, R. Beare, V. Srikanth, and H. Ma, "Googling location for Mobile Stroke Unit hub in metropolitan Sydney," *Frontiers in neurology*, vol. 10, p. 810, 2019.

[14] J. P. Rhudy Jr, A. W. Alexandrov, J. Rike, T. Bryndziar, A. H. Z. Maleki, V. Swatzell, W. Dusenbury, E. J. Metter, and A. V. Alexandrov, "Geospatial visualization of mobile stroke unit dispatches: a method to optimize service performance," *Interventional neurology*, vol. 7, no. 6, pp. 464-470, 2018.

[15] S. Mathur, S. Walter, I. Q. Grunwald, S. A. Helwig, M. Lesmeister, and K. Fassbender, "Improving prehospital stroke services in rural and underserved settings with mobile stroke units," *Frontiers in Neurology*, vol. 10, 2019.

[16] Statistics Sweden, "demographic data 2018," [Online]. Available: https://www.scb.se. [Accessed: Sept. 6, 2018].

[17] S. C. Ajmi, R. Advani, L. Fjetland, K. D. Kurz, T. Lindner, S. A. Qvindesland, H. Ersdal, M. Goyal, J. T. Kvaløy, and M. Kurz, "Reducing door-to-needle times in stroke thrombolysis to 13 min through protocol revision and simulation training: a quality improvement project in a Norwegian stroke centre," *BMJ quality & safety*, vol. 28, no. 11, pp. 939-948, 2019.

[18] Sweden's Southern Regional Health Care Committee, "stroke data 2018," [Online]. Available: https://sodrasjukvardsregionen.se. [Accessed: Sept. 6, 2018].

PAPER II - AN OPTIMIZATION MODEL FOR THE TRADEOFF BETWEEN EFFICIENCY AND EQUITY FOR MOBILE STROKE UNIT PLACEMENT

Saeid Amouzad Mahdiraji, Johan Holmgren, Radu-Casian Mihailescu, and Jesper Petersson

ABSTRACT

A mobile stroke unit (MSU) is an ambulance, where stroke patients can be diagnosed and treated. Recently, placement of MSUs has been studied focusing on either maximum population coverage or equal service for all patients, termed efficiency and equity, respectively. In this study, we propose an unconstrained optimization model for the placement of MSUs, designed to introduce a tradeoff between efficiency and equity. The tradeoff is based on the concepts of weighted average time to treatment and the time difference between the expected time to treatment for different geographical areas. We conduct a case-study for Sweden's Southern Health care Region (SHR), generating three scenarios (MSU1, MSU2, and MSU3) including 1, 2, and 3 MSUs, respectively. We show that our proposed optimization model can tune the tradeoff between the efficiency and equity perspectives for the MSU(s) allocation. This enables a high level of equal service for most inhabitants, as well as reducing the time to treatment for most inhabitants of a geographic region. In particular, placing three MSUs in the SHR with the proposed tradeoff, the share of inhabitants who are expected to receive treatment within an hour potentially improved by about a factor of 14 in our model.

PUBLISHED IN

9th KES International Conference on Innovation in Medicine and Healthcare (KES-InMed-21), pp. 183-193. Springer, Singapore, 2021.

1. Introduction

Stroke, which is a medical condition resulting in reduced blood flow in the brain, is the second most common cause of death worldwide and a leading cause of permanent physical and cognitive disability, leaving people paralyzed and unable to perform their daily activities [1]. There are three main types of stroke: ischemic, hemorrhagic, and transient ischemic attack (TIA), each requiring specific treatment. In ischemic stroke, one or more clots reduce the blood flow inside the brain, and patients should receive thrombolysis and sometimes thrombectomy if the clot is large. A hemorrhagic stroke occurs when a blood vessel in the brain ruptures and blood flows into the surrounding tissues. Current guidelines for hemorrhagic stroke recommend very early start of blood pressure lowering therapy. A TIA occurs when the blood flow is temporarily blocked during a short period of time, allowing the brain function to fully recover. The type of stroke is determined by performing a computed tomography (CT) scan on the brain of the patient.

It is well established that immediate treatment is essential for all types of stroke; in particular, the term “golden hour” is proposed for ischemic stroke, asserting that the patients who receive treatment within an hour of symptom onset, have a much higher chance for full recovery than the patients with later start of treatment [2]. However, due to logistical challenges, it is often difficult to provide fast enough treatment for stroke patients. As CT scanners are typically only available at hospitals, the patient first must be transported to an acute hospital for diagnosis and treatment.

Mobile stroke units (MSUs), in addition to regular ambulances, have been deployed in some areas (for example: Berlin, Cleveland and Melbourne) as an alternative for prehospital diagnosis and care [2]. An MSU is a specialized ambulance equipped with a CT scanner allowing the ambulance personnel to diagnose the stroke patients and provide intravenous stroke treatment in the ambulance. Therefore, the use of MSUs reduces at least the time required to transport and to diagnosis the patient at the hospital.

However, MSUs are expensive, and it is important to locate them in order to provide maximum benefit for the patients. Optimization modeling has been used to identify efficient ambulance/MSU locations in a region using two perspectives: efficiency and equity [3-5]. The majority of the existing studies aim to reach maximum population coverage by considering efficiency, i.e., placing ambulances in an optimal way to cover as many persons as possible. Equity, which is covered in some studies, seeks to provide equal service for all patients, regardless of where they live. It should be emphasized that each of these perspectives has a bias towards a group of inhabitants, i.e., residents of densely populated or rural areas.

We contribute an objective function used in optimization modeling in order to tune the tradeoff between efficiency and equity for the optimal placement of MSUs in a geographic region. In a scenario study, we evaluate the proposed model by comparing the current situation in the Sweden's Southern Health care Region (SHR) with three generated MSU scenarios. The computational results show that the presented approach can be used to tune the tradeoff between efficiency and equity.

The rest of this paper is organized as follows. Section 2 reviews the related work. Section 3 introduces the time to treatment estimation model. Section 4 presents an optimization model to make a tradeoff between efficiency and equity for optimal MSU(s) placement. Our scenario study is presented is Section 5 and the results and discussion are given in Section 6. Finally, Section 7 concludes the paper.

2. Related Work

In the emergency medical services (EMS) literature, efficiency and equity have been assessed by the coverage of urban and rural areas, respectively. Likewise, efficiency emphasizes the deployment of an MSU in a place where it potentially helps a higher amount of people to get a shorter time to thrombolysis. The focus of equity is to cut down the time to thrombolysis for patients residing far from the hospitals providing thrombolysis and CT scan. However, while the goal of efficiency is to provide maximum population coverage, there is no agreed definition of equity or how it should be measured. Equitable service could be provided using range [6], variance [7], mean absolute deviation [8], squared coefficient of variation [9], Gini coefficient [10], and envy criteria [11].

The tradeoff between efficiency and equity for optimal placement of regular ambulances has recently gained attention in the research community. Enayati et al. [10] use multi-objective optimization for location and dispatching problems and achieve a balanced solution considering both efficiency and equity. Chanta et al. [12] use a bicriteria optimization framework to combine efficiency and equity. Toro-Díaz et al. [9] develop a large-scale EMS system by considering the efficiency and equity criteria using a Tabu Search-based heuristic with an embedded approximation procedure.

Some articles have been written addressing optimal placement of MSUs, with main focus on the benefits for residents in urban areas (efficiency perspective) [3, 5] or rural areas (equity perspective) [4]. Rhudy Jr. et al. [5] employ a geospatial analysis of the distribution of MSUs to optimize service delivery for stroke patients in the city of Memphis. Phan et al. [3] use Google maps to find the optimal location of an MSU in Sydney. Dahllöf et al. [13] use expected value optimization in order to identify the optimal placement of an MSU in the Skåne county of Sweden.

To the best of our knowledge, no prior studies have tried to optimally place an MSU in a geographical region considering a tradeoff between efficiency and equity. In the current study, we extend the work of Dahllöf et al. [13] by explicitly incorporating the tradeoff between efficiency and equity. While the abovementioned studies have a restricted set-up with only one MSU, in the present paper, we provide a generalized analysis for one or more MSUs. Finally, while prior studies mostly assess their approaches in highly populated areas, we apply our approach in Sweden's Southern Health care Region, which includes both urban and rural areas.

3. Time to Treatment Estimation Model

In this section, we present our time to treatment estimation model for the patients located at different places in a geographical region. See the companion paper by Amouzad Mahdiraji et al. [14], for a more detailed description of the model.

We divided the region of study (ROI) into disjoint sub-regions, a set of 1×1 km squares R, enabling us to take into account the variation of population density and the expected time to treatment over various parts of the ROI. In our calculations, we assumed that all of the patients in the square $r \in R$ are located in its center c_r, which means that all transport to and from sub-region $r \in R$, is assumed to be made to and from the center c_r (of r).

Let L^{AMB}, L^{MSU}, and L^{H} denote the set of regular ambulance sites, MSU locations, and acute hospital locations in the ROI, respectively. For a regular ambulance located at ambulance site $l \in L^{AMB}$, we let $t_l^{AMB\ RESP}$ be the expected response time, i.e., the time from an emergency call until an ambulance dispatches, $t_{lr}^{AMB\ LR}$ the expected time to drive from ambulance location l to the centroid c_r of sub-region r, $t^{AMB\ LAY}$ the expected layover time, i.e., the time from the ambulance has arrived at the patient location until it departs, $t_{rh}^{AMB\ RH}$ the expected time to drive from c_r to acute hospital $h \in L^{H}$, and t_h^{DTN} the expected time for diagnosis at acute hospital $h \in L^H$, i.e., the expected time from the arrival of the patient to the hospital until the treatment is initiated.

For an MSU located at ambulance site $l \in L^{MSU}$, $t_l^{MSU\ RESP}$ is the expected time from an emergency call until an MSU starts driving towards the patient site, $t_{lr}^{MSU\ LR}$ is the expected time to drive from l to c_r, $t^{MSU\ LAY}$ is the expected layover time for an MSU, and $t^{MSU\ DIAG}$ is the expected time to diagnose a stroke patient inside an MSU.

The expected time to treatment, using only the regular ambulances located in L^{AMB}, for a patient located in square $r \in R$ is estimated as:

$$t_r^{AMB\ TT} = \min_{l \in L, h \in H} \{t_l^{AMB\ RESP} + t_{lr}^{AMB\ LR} + t^{AMB\ LAY} + t_{rh}^{AMB\ RH} + t_h^{DTN}\}. \tag{1}$$

Equation (1) is the minimum expected time to treatment with respect to the nearest ambulance site l and the nearest acute hospital h to the patient located in square r.

Assuming that only the MSUs located in L^{MSU} can be used, the expected time to treatment for a patient located in square $r \in R$ is estimated as:

$$t_r^{MSU\ TT} = \min_{l \in L, h \in H} \{t_l^{MSU\ RESP} + t_{lr}^{MSU\ LR} + t^{MSU\ LAY} + t^{MSU\ DIAG}\}. \tag{2}$$

When both regular ambulances and MSUs are available, the expected time to treatment for a patient located in square $r \in R$ is estimated as:

$$t_r^{TT} = \min \{t_r^{AMB\ TT}, t_r^{MSU\ TT}\}. \tag{3}$$

4. Tradeoff Between the Efficiency and Equity Perspectives

In the health care domain, efficiency can be loosely defined as helping as many patients as possible as fast as possible. Efficiency can be measured in terms of the total cost for society or the average time to treatment for the patients with the same diagnosis. In the current study, we chose the time to treatment as efficiency measure; in particular we used the weighted average time to treatment (WATT) t^{TT}, which is calculated as:

$$t^{TT} = \sum_{r \in R} t_r^{TT} \cdot Q_r, \tag{4}$$

where Q_r is the share of stroke cases (in the ROI) that is expected to take place in sub-region $r \in R$ ($\sum_{r \in R} Q_r = 1$) and t_r^{TT} is the expected time to treatment for a patient located in sub-region $r \in R$.

Efficiency aims to place MSUs where they reduce the expected time to treatment most for the whole region. Urban areas seem to be a favorable location of MSUs in this perspective due to being highly populated. The focus of our choice of efficiency definition is to find an MSU location set that minimizes equation (4). This corresponds to the optimization problem:

$$\underset{s \in S}{argmin}\ t_s^{TT} = \left\{\sum_{r \in R} t_{r,s}^{TT} \cdot Q_r\right\}, \tag{5}$$

where S denotes the set of all possible MSU allocations, or in other words, all possible solutions to the optimization problem, and $s \in S$ denotes a particular MSU allocation. It is worth noting that s defines whether or not an MSU is allocated to each of the ambulance locations. Furthermore, t_s^{TT} is the WATT for the whole ROI considering the MSU allocation s and $t_{r,s}^{TT}$ is the expected time to treatment for sub-region $r \in R$ considering ambulance allocation s.

In a system with perfect equity, all patients have the same time to treatment, regardless of who they are and where they are located. Several quantitative measures have been suggested in order to measure the equity of a system. Among the equity measures described in Section 2, we believe that the range, variance, mean absolute deviation, squared coefficient of variation, and Gini index are the most practical for the MSU location problem as they are among the strictest measures of equity and they do not have a built-in tradeoff due to variation of the population density. We made a comparison between the mentioned equity measures and decided to employ the range as our measure of equity. The range measure strives to minimize the difference between the expected time to treatment for patients located at different places in the ROI. The optimal location, considering the equity perspective, of MSUs can be identified by solving the following optimization model:

$$\underset{s \in S}{argmin}\ \left\{max\{t_{r,s}^{TT}\} - min\{t_{r,s}^{TT}\}\right\}. \tag{6}$$

We make use of the efficiency and equity equations defined above to establish our tradeoff function, which is the objective function in the optimization problem:

$$\underset{s \in S}{argmin} \{(1 - W)(F_{efficiency}) + W(G_{equity})\}, \tag{7}$$

where $F_{efficiency}$ and G_{equity} are the efficiency (see equation (5)) and equity (see equation (6)) functions, respectively. W is a controlling weight, which is used to determine the impact of each of the efficiency and equity on the allocation of MSUs, varying from 0 to 1. After assigning a value to W, the optimization problem is solved by finding the MSU allocation s that minimizes the equation (7). In particular, when $W = 0$ or $W = 1$, only efficiency or equity is considered, respectively. If, for example, $W = 0.5$, equal weight is given to efficiency and equity. The reason that we applied the weight W to the tradeoff function is to enable to support the public health authorities' priorities regarding efficiency and equity when deciding where to place the MSU(s).

For smaller problem instances, our optimization problem can be solved using exhaustive search; however, larger problem instances might require more sophisticated solution methods. It is also possible to add additional constrains to the presented optimization models, e.g., limitation of the number of MSUs to locate, which is a restriction of the solution set (i.e., S).

5. Scenario Study

In order to evaluate the proposed optimization model for trading off between efficiency and equity when placing MSUs, we have applied our optimization model (see equation (7)) on Sweden's Southern Health care Region (SHR), where about 3900 stroke incidents occur annually [15]. The SHR consists of 4 counties and 49 municipalities, where there are 13 acute hospitals equipped with CT scanner and 39 ambulance sites. An overview of the SHR is provided in Figure 1, where the green triangles and purple circles represent the locations of ambulance sites and acute hospitals, respectively. For future reference, each circled number in Figure 1 specifies the corresponding ambulance site ID.

We used two types of data: demographic data from Statistics Sweden [15] and stroke data for 2018, provided by Sweden's Southern Regional Health Care Committee [16]. The ROI (i.e., SHR) was divided into a disjoint set of 1×1 km squares denoted by $r \in R$. The union of all squares $\bigcup_{r \in R} r$ equals to the SHR. The location of ambulance sites and acute hospitals were acquired using Google maps and official documentation provided by the health care authorities in the region.

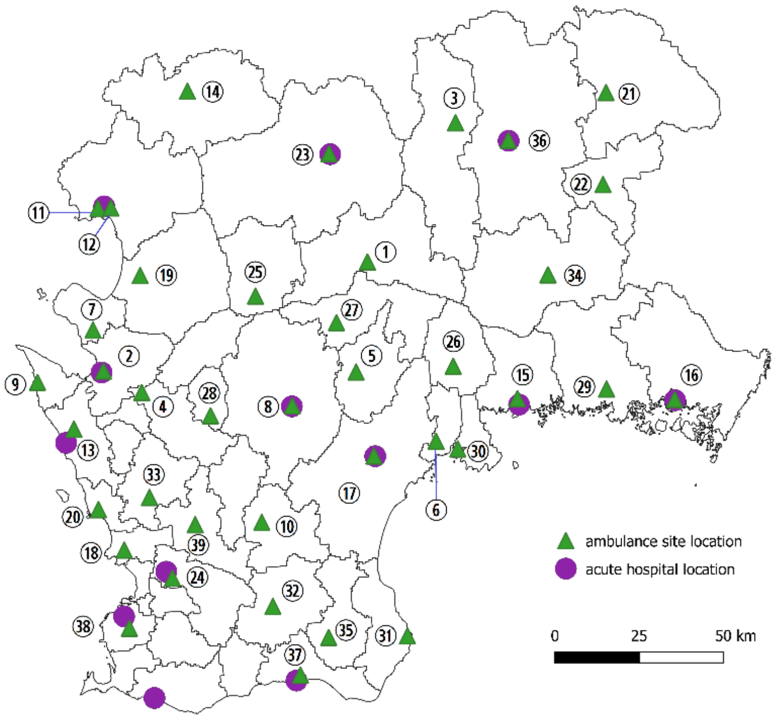


Figure 1. An overview of SHR, where ambulance sites and acute hospital locations are shown by green triangles and purple circles, respectively. The circled numbers show the corresponding ambulance site ID.

The stroke data included the number of stroke cases for 21 age groups, $\{[0,4), [4,8), \ldots, [95,99), [100, \infty)\}$. In addition, the demographic data contained the number of inhabitants for each age-group and each of the 1×1 km squares covering the SHR. Using this information, we calculated, for each age-group, the likelihood that a person will get a stroke during a specific time-period, i.e., the year 2018 in this study. Using the calculated stroke likelihoods for each sub-region $r \in R$, we calculated the expected number of stroke incidents I_r for each of the sub-regions $r \in R$. The share of the stroke incidents that is expected to occur in sub-region $r \in R$ is given by:

$$Q_r = \frac{I_r}{I}, \tag{8}$$

where $I = \sum_{r \in R} I_r$ is the expected number of stroke cases in SHR.

It should be emphasized that efficiency and equity may have different range of values; hence, it is likely that one of them dominates the other in the tradeoff function (equation (7)). In order to tackle this problem, we normalized the values of each perspective, using min-max scaling, before applying them to equation (7).

In our scenario study, we created four different scenarios for possible MSU allocations. The first one is a baseline scenario, the current situation where only regular ambulances are located in all 39 ambulance sites in the SHR. We also created three scenarios including 1 (called MSU1), 2 (called MSU2), and 3 MSUs (called MSU3), respectively, in addition to the regular ambulances. In MSU1-MSU3, we assumed that all 39 ambulance sites located in the SHR are candidate locations for MSUs.

For example, in MSU2, the MSU location set S contains all possible ways to allocate two MSU2, i.e., the following combinations of ambulance locations $\{[1,2],[1,3],\ldots,[1,39],\ldots,[38,39]$. For example, $[1,3]$ means that MSUs are located in the ambulance sites 1 and 3 (see Figure 1). For each of the considered scenarios (baseline, MSU1, MSU2, and MSU3), we solved our tradeoff optimization problem (equation (7)) for each of a number of tradeoff weights $W = \{0, 0.05, 0.1, \ldots, 0.95, 1\}$.

It should be noted that each optimization problem contains 2^n solutions (i.e., possible ambulance allocations S, where $n = |L^{MSU}|$). In order to solve the optimization problem, we enumerated over all of the possible solutions.

We estimated the driving times of both regular ambulances and MSUs using the driving time generation functionality provided by the open street map (OSM).

Due to the limitations of the available data, we consulted a neurologist with insight into stroke logistics in order to make a few other assumptions. We assumed that a regular ambulance drives 5% faster than a normal car, and an MSU drives at the same speed as a normal car; the response time for both regular ambulances and MSUs is 3 minutes, $t_l^{AMB\ RESP} = t_l^{MSU\ RESP} = 0.05\text{h}$; the layover time for both regular ambulances and MSUs is 15 minutes, i.e., $t^{AMB\ LAY} = t^{MSU\ LAY} = 0.25\text{h}$; the time for diagnosis of a patient inside an MSU is 15 minutes, $t^{MSU\ DIAG} = 0.25\text{h}$; the expected time for diagnosis for each of the considered hospitals is 35 minutes, $t_h^{DTN} = 0.583\text{h}$; it is assumed that ambulance sites and hospitals are open and can provide 24/7 service; the MSUs in each MSU scenario can provide service over the whole SHR; and all stroke cases occur at the home of the patients.

6. Results and Discussion

In Figure 2, we display, for each of our MSU scenarios, the performance of the tradeoff function regarding the WATT and the range when solving the optimization model for different values of W. In each of the plots, the x-axis is the W, which varies between 0 and 1, and the y-axis is the normalized value. When $W = 0$, the tradeoff represents the efficiency perspective; however, as W increases towards 1, the tradeoff is gradually shifted towards equity.

In Table 1, we compare the performance of the efficiency, equity, and tradeoff (when efficiency and equity have the same weight in the tradeoff, $W = 0.5$) for each of the three MSU scenarios with the baseline scenario. The results indicate that after placing MSUs in the SHR, the WATT will decrease for all MSU scenarios, and the higher the number of MSUs is, the more the reduction in the WATT is expected. However, the tradeoff strives to significantly shorten not only the WATT but also the range for all of the three MSU scenarios in comparison to the baseline. In particular, the results of the tradeoff, when the weight is 0.5, reduced the WATT by 3, 13.2, and 16.8 minutes respectively for MSU1, MSU2, and MSU3. Furthermore, the range measure is expected to be reduced from 1.91 hours for the baseline to 1.67 hours for MSU1 and MSU2; and 1.6 hours for MSU3.

a) MSU1

b) MSU2

b) MSU3

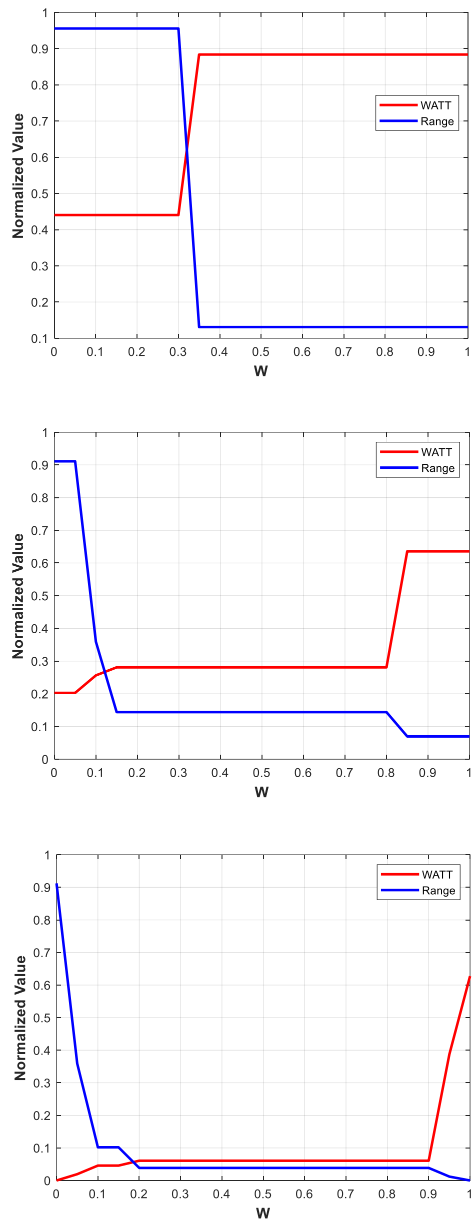


Figure 2. Performance of the tradeoff the WATT and the range for different values of W concerning a) MSU1, b) MSU2, and c) MSU3.

In Table 1, we also present the share of the total population of SHR, i.e., 1687190, that are expected to get treatment within 60, 75, and 90 minutes for each of the scenarios and perspectives. The share of the population who are expected to receive treatment within 60 minutes is expected to increase from 3.96% for the baseline to 9.2%, 45.37%, and 55.02% for MSU1, MSU2, and MSU3, respectively. In particular, for MSU3, placing 3 MSUs in the SHR, using the proposed tradeoff, is expected to lead to an improvement of approximately a factor 14 concerning the share of persons who get treatment within an hour. It is noteworthy that comparing the results for the different perspectives verify that the WATT has an inverse relationship with the share of persons who are expected to receive treatment within 60 minutes.

The results also indicate that after placing MSU(s) using the tradeoff weight $W = 0.5$, the time to treatment is expected to decrease for 29% (187150 inhabitants), 59% (1003167 inhabitants), and 75% (1272676 inhabitants) of the total population considering MSU1, MSU2, and MSU3, respectively. In particular, the time to treatment of 142604, 799336, and 1023684 inhabitants is expected to decrease up to 30 minutes.

Table 1. WATT (in hours), range (in hours), and the share of the inhabitants in the SHR who are expected to receive treatment within 60, 75, and 90 minutes for each scenario and perspective. The information within square brackets is the optimal ambulance site IDs.

Perspective	Scenario	MSU sites	WATT t^{TT}	Range	Expected time to treatment		
					< 60 minutes	< 75 minutes	< 90 minutes
	Baseline	-	1.33	1.91	3.96%	47.76%	80.89%
Efficiency	MSU1	[38]	1.16	2.24	40.13%	66.76%	86.59%
	MSU2	[13, 38]	1.09	2.21	51.54%	74.71%	89.55%
	MSU3	[13,17,38]	1.04	2.21	57.67%	81.25%	92.44%
Equity	MSU1	[3]	1.28	1.67	9.2%	51.37%	84.43%
	MSU2	[2, 3]	1.21	1.62	18.85%	61.86%	88.44%
	MSU3	[10,14, 22]	1.21	1.57	12.24%	64.33%	91.13%
Tradeoff [0.5, 0.5]	MSU1	[3]	1.28	1.67	9.2%	51.37%	84.43%
	MSU2	[3, 38]	1.11	1.67	45.37%	70.38%	90.13%
	MSU3	[2,3, 38]	1.05	1.60	55.02%	78.52%	93.21%

7. Conclusions

We contribute an unconstrained optimization algorithm designed to introduce a tradeoff between the two perspectives of efficiency and equity when allocating mobile stroke units (MSUs) in a geographical region. The weighted average time to treatment (WATT) and range were our choice of efficiency and equity measures, respectively. In a scenario study, we evaluated our optimization model by comparing the current situation, represented by a baseline scenario, in the Sweden's Southern Health care Region (SHR) with three MSU scenarios including 1, 2, and 3 MSUs, respectively. Our experimental results show that the use of the proposed tradeoff function has the potential to balance efficiency and equity by reducing both WATT and range compared to the baseline scenario. We also conclude that the use of the proposed optimization model for MSU placement may contribute to equal service for most inhabitants as well as substantially increasing the number of persons who are expected to get treatment within 60 minutes.

References

[1] World Stroke Organization, "Facts and figures about stroke," [Online]. Available: https://www.world-stroke.org/world-stroke-day-campaign/whystroke-matters/learn-about-stroke. [Accessed: Dec. 20, 2019].

[2] M. Ebinger, B. Winter, M. Wendt, J. E. Weber, C. Waldschmidt, M. Rozanski, A. Kunz, P. Koch, P. A. Kellner, and D. Gierhake, "Effect of the use of ambulance-based thrombolysis on time to thrombolysis in acute ischemic stroke: a randomized clinical trial," *Jama*, vol. 311, no. 16, pp. 1622-1631, 2014.

[3] T. G. Phan, R. Beare, V. Srikanth, and H. Ma, "Googling location for Mobile Stroke Unit hub in metropolitan Sydney," *Frontiers in neurology*, vol. 10, p. 810, 2019.

[4] S. Mathur, S. Walter, I. Q. Grunwald, S. A. Helwig, M. Lesmeister, and K. Fassbender, "Improving prehospital stroke services in rural and underserved settings with mobile stroke units," *Frontiers in Neurology*, vol. 10, 2019.

[5] J. P. Rhudy Jr, A. W. Alexandrov, J. Rike, T. Bryndziar, A. H. Z. Maleki, V. Swatzell, W. Dusenbury, E. J. Metter, and A. V. Alexandrov, "Geospatial visualization of mobile stroke unit dispatches: a method to optimize service performance," *Interventional neurology*, vol. 7, no. 6, pp. 464-470, 2018.

[6] T. Drezner and Z. Drezner, "Equity models in planar location," *Computational Management Science*, vol. 4, no. 1, pp. 1-16, 2007.

[7] T. Drezner and Z. Drezner, "A note on equity across groups in facility location," *Naval Research Logistics (NRL)*, vol. 58, no. 7, pp. 705-711, 2011.

[8] G. F. Mulligan, "Equality measures and facility location," *Papers in Regional Science*, vol. 70, no. 4, pp. 345-365, 1991.

[9] H. Toro-Díaz, M. E. Mayorga, L. A. McLay, H. K. Rajagopalan, and C. Saydam, "Reducing disparities in large-scale emergency medical service systems," *Journal of the Operational Research Society*, vol. 66, no. 7, pp. 1169-1181, 2015.

[10] S. Enayati, M. E. Mayorga, H. Toro-Díaz, and L. A. Albert, "Identifying trade-offs in equity and efficiency for simultaneously optimizing location and multipriority dispatch of ambulances," *International Transactions in Operational Research*, vol. 26, no. 2, pp. 415-438, 2019.

[11] Espejo, A. Marín, J. Puerto, and A. M. Rodríguez-Chía, "A comparison of formulations and solution methods for the minimum-envy location problem," *Computers & Operations Research*, vol. 36, no. 6, pp. 1966-1981, 2009.

[12] S. Chanta, M. E. Mayorga, and L. A. McLay, “Improving emergency service in rural areas: a bi-objective covering location model for EMS systems,” *Annals of Operations Research*, vol. 221, no. 1, pp. 133-159, 2014.

[13] O. Dahllöf, F. Hofwimmer, J. Holmgren, and J. Petersson, “Optimal placement of Mobile Stroke Units considering the perspectives of equality and efficiency,” in *Proceedings of the 8th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH 2018)*, vol. 141, pp. 311-318, 2018.

[14] S. Amouzad Mahdiraji, O. Dahllöf, F. Hofwimmer, J. Holmgren, R.-C. Mihailescu, and J. Petersson, “Mobile stroke units for acute stroke care in the south of Sweden,” Cogent Engineering, vol. 8, no. 1, 2021, doi: https://doi.org/10.1080/23311916.2021.1874084.

[15] Statistics Sweden, “demographic data 2018,” [Online]. Available: https://www.scb.se. [Accessed: Sept. 6, 2018].

[16] Sweden’s Southern Regional Health Care Committee, “stroke data 2018,” [Online]. Available: https://sodrasjukvardsregionen.se. [Accessed: Sept. 6, 2018].

PAPER III - A MICRO-LEVEL SIMULATION MODEL FOR ANALYZING THE USE OF MSUS IN SOUTHERN SWEDEN

Saeid Amouzad Mahdiraji, Johan Holmgren, Radu-Casian Mihailescu, and Jesper Petersson

ABSTRACT

A mobile stroke unit (MSU) is a special type of ambulance, where stroke patients can be diagnosed and provided intravenous treatment, hence allowing to cut down the time to treatment for stroke patients. We present a discrete event simulation (DES) model to study the potential benefits of using MSUs in the southern health care region of Sweden (SHR). We included the activities and actions used in the SHR for stroke patient transportation as events in the DES model, and we generated a synthetic set of stroke patients as input for the simulation model. In a scenario study, we compared two scenarios, including three MSUs each, with the current situation, having only regular ambulances. We also performed a sensitivity analysis to further evaluate the presented DES model. For both MSU scenarios, our simulation results indicate that the average time to treatment is expected to decrease for the whole region and for each municipality of SHR. For example, the average time to treatment in the SHR is reduced from 1.31h in the baseline scenario to 1.20h and 1.23h for the two MSU scenarios. In addition, the share of stroke patients who are expected to receive treatment within one hour is increased by a factor of about 3 for both MSU scenarios.

PUBLISHED IN

Proceedings of the 11th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH 2021), vol. 198, pp. 132-139, 2022.

1. Introduction

Stroke is a leading cause of death and disability around the world [1]. A stroke occurs when a blood clot or bleeding in the brain reduces the blood flow to the brain. An ischemic stroke, which is the most common stroke type, happens when one or more clots reduce the blood flow inside the brain, and patients should receive thrombolysis and sometimes thrombectomy, depending on the size of the clot. The type of stroke can be determined by a computed tomography (CT) scan on the brain of the patient.

Treatment should be instigated as early as possible. In particular, a 60-minute period known as the “golden hour” has been introduced in ischemic stroke treatment [2] to emphasize the importance of early treatment. However, due to logistical challenges, it is often difficult to provide rapid treatment for stroke patients. Since CT scanners are in general only available at hospitals, it is not possible to start diagnosis and treatment until after the patient has been transported to the nearest acute hospital.

A mobile stroke unit (MSU) is a special stroke ambulance that has become an alternative for prehospital diagnosis and treatment of stroke patients [2]. An MSU is equipped with a CT scanner, enabling ambulance personnel to diagnose the stroke patients and provide intravenous stroke treatment already inside the ambulance. Thus, the use of MSUs, as the MSU operations in Berlin, Cleveland, and Melbourne suggest, has the potential to cut down the time to treatment for many stroke patients [2-4].

To improve the stroke transport logistics, different policies, including the use of MSUs, can be implemented. However, before implementing a policy, its performance needs to be assessed. Due to the difficulties of evaluating decision policies using real patients, the use of simulation is preferable. Simulation can be used to assess policies before being implemented in the real system while taking into account various attributes of the stroke population, such as population size, stroke time, and patient locations. Simulation also enables to analyze policies without risking the health of the patients who are already in a vulnerable condition. The simulation output can be used by decision-makers when deciding which policy to apply in a particular region.

Agent-based simulation and discrete event simulation (DES) have been used for the analysis of emergency medical service (EMS) transport in stroke. For example, Bogle et al. [5] present a DES model to evaluate the effects of altering important specifications of an EMS routing algorithm for patients with large vessel occlusion stroke in two different counties in the USA. Al Fatah et al. [6]

use agent-based simulation to assess two stroke transport policies regarding where to transport suspected stroke patients for diagnosis, i.e., *nearest hospital* policy and *nearest hospital towards the stroke center* policy, in the southern healthcare region of Sweden (SHR). However, the mentioned studies only consider regular ambulances, and to the best of our knowledge, no previous studies use simulation to evaluate MSU related transport policies.

In a previous study [7], we propose a macro-level average time to treatment estimation model in order to analyze the potential benefits of placing MSUs in the SHR. The previously proposed model is not able to study the individual patients and the individual emergency vehicles (EVs) as it generates an average expected time to treatment for a whole population. In addition, the macro-level modeling paradigm is unable to take into account the effects of simultaneous stroke incidents, e.g., if an MSU is needed at two places at the same time. In the current paper, we present a DES model for evaluating prehospital stroke transport policies related to the use of MSUs. This type of modeling allows to simulate the activities of individual entities over time; hence, allowing more realistic modeling. In addition, it enables to add stochasticity, including location and time of stroke incidents, to the model. Thus, the proposed model is able to simulate the main actions and decisions involved in the logistical operations of stroke patients, and it takes as input a population of stroke patients. In a case study, set in the SHR, we use our simulation model to compare a baseline scenario containing only regular ambulances with two MSU scenarios, each containing three MSUs.

The rest of this paper is organized as follows. Section 2 introduces our DES model for MSU-related stroke logistics policies assessment. Section 3 presents the scenario study, followed by the results and discussion. Finally, Section 4 concludes the paper.

2. Discrete Event Simulation Model

In this section, we describe our simulation model for the evaluation of MSU-related stroke logistics policies. The model makes use of a zone-based approach, and it simulates the actions applied for each individual patient. The geographical region considered by the model is divided into a non-overlapping set of subregions, where it is assumed that all of the patients in each subregion are located in its centroid, and all transports to and from a specific subregion are made to and from its centroid.

The simulation model consists of three main components: input, simulation model, and output. The *input* is required in order to run the simulation model and to regulate everything that is not static. As input, the model takes a synthetic stroke population, data about hospitals and EVs, and transport data, including the driving times between EV stations and patient locations, and between patient locations and hospitals. The *simulation model* includes all of the logic, and it generates a log of all activities that have occurred as *output*, including the start and end time for each of the simulated activities. The total service time of EVs and the time to treatment for each patient are examples of output variables.

Our model is implemented based on the principle of the DES framework, which is a paradigm used to model the operation of a system as a series of events that occur over time, where an event is an instantaneous occurrence that may change the state of the system. In our model, each action is modeled using two events, i.e., a start event and an end event, enabling to model that the simulated actions do not occur instantaneously. Figure 1 represents the overall flow of our simulation model. The model has a clock representing the simulated time and a queue of future events that are scheduled for processing. The size of the queue is not static as the events are dynamically added and removed at run-time. During the simulation, new events are continuously created and sorted into the queue based on their time of occurrence; the next event is always the first event in the queue. The occurrence of a certain event causes an action or a set of actions, and often, adding a new event to the queue. Each of the events will occur at a certain point of time during the simulation, where it represents either a starting action and/or ending action. At the end of each event, the corresponding event will be removed from the event queue, and the simulation clock will advance to the time of the next event in the queue. Once event queue is empty, or the simulation end time has been reached, the simulation terminates.

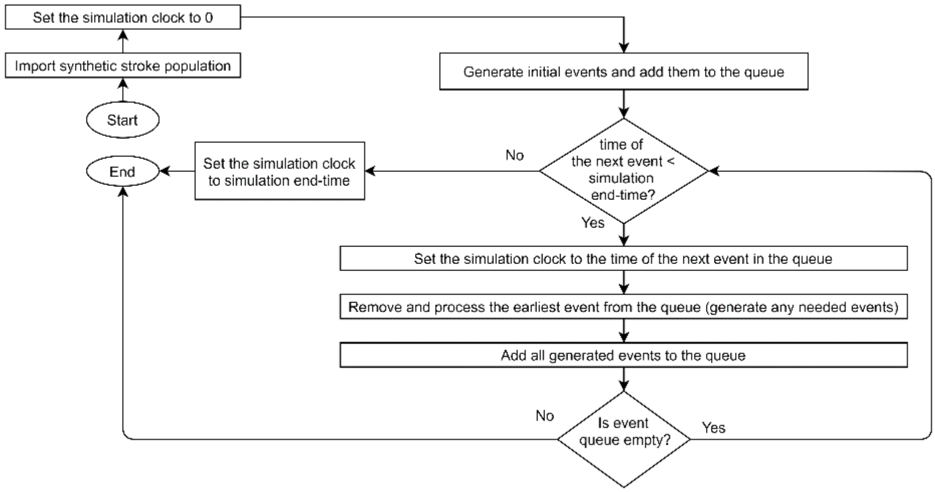


Figure 1. Flowchart of the DES algorithm used in our simulation model.

For each individual in a set of stroke patients, the model simulates the main actions and decision-making that occur from the time of receiving a call concerning a suspected stroke incident until the thrombolysis is initiated. See Figure 2 for the chain of actions that are expected to take place for each stroke patient in our DES model. The actions and decisions included in the model are based on the current stroke logistics process for an ischemic stroke patient requiring thrombolysis adopted in the SHR. The chain of activities for a particular patient is initiated by a stroke incident, which triggers a call to the emergency center. The subsequent actions will be created for each patient during the simulation according to the care chain presented in Figure 2. For patients transported by a regular ambulance, the time to treatment is the expected time from when a stroke happens until the patient receives thrombolysis in the acute hospital. For the MSU cases, the time to treatment is the expected time from when a stroke happens until the patient receives thrombolysis inside the MSU. When an EV arrives at the acute hospital and drops off the patient, it becomes available for the next patient. Please note that we assumed that all the modeled patients are suspected stroke patients, and that they need to be transported to an acute hospital.

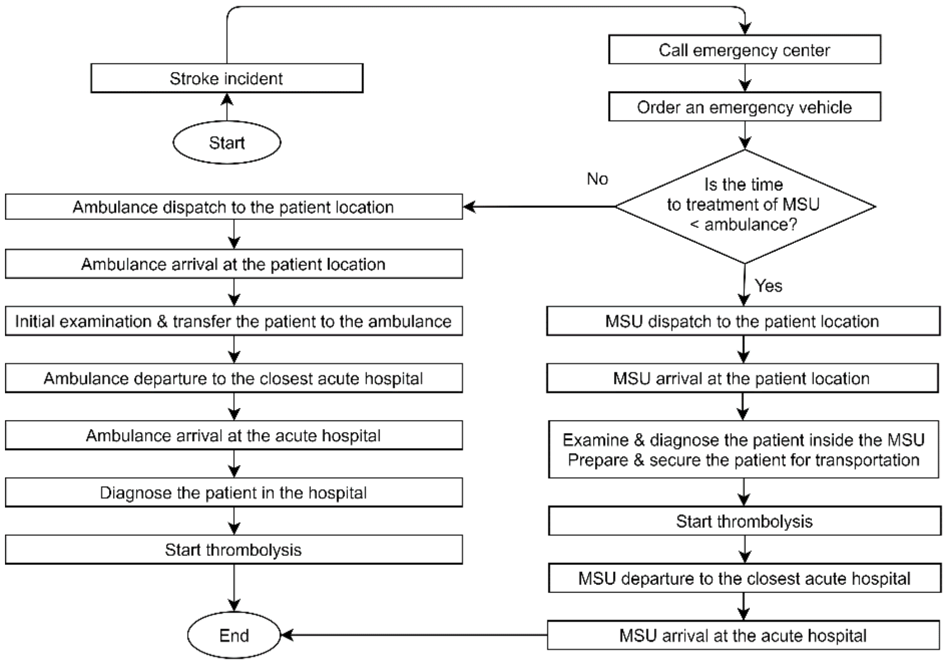


Figure 2. Flowchart demonstrating the sequence of actions that happen for each stroke patient during the simulation. Please note that the term ambulance is used instead of regular ambulance in the flowchart.

Once a stroke incident occurs, the model compares the expected travel times of different available EVs between the EV sites and the patient location and then dispatches the EV, either a regular ambulance or an MSU, to the scene, depending on which vehicle is expected to provide the fastest time to treatment. When there is no available regular ambulance or MSU sufficiently close for a stroke case, the model estimates which busy EV can provide the fastest time to treatment and then waits until the chosen busy EV becomes available. In such a situation, there would be waiting time, which is defined for each stroke event as the time difference between when the chosen EV becomes available and the post-response time for the current stroke event.

3. Scenario Study

In a scenario study, we applied our DES model to study the potential effects of using MSUs in Sweden's Southern Swedish Health care Region (SHR), where about 3900 individuals suffer a stroke each year [8]. The SHR includes 4 counties and 49 municipalities, and it contains 13 acute hospitals equipped with CT scanner and 39 ambulance sites (see Figure 3 (a)).

We used two types of data, i.e., demographic data for 2018 from Statistics Sweden [9] and stroke data from Sweden's Southern Regional Health Care Committee [10]. The demographic data includes the number of residents for 21 age-groups, i.e., $\{[0,4), [4,8), \dots, [95,99), [100, \infty)\}$, for each of the set of disjoint subregions of SHR. The stroke data set consists of the number of stroke cases for each municipality of SHR and each age-group, as well as the aggregated times of strokes for 2016.

In our scenario study, we considered three different scenarios: a baseline scenario and 2 MSU scenarios with different characteristics. The baseline scenario corresponds to the current situation in the SHR, where only regular ambulances are used. The two MSU scenarios are denoted as MSU1 and MSU2, and each of them contains 3 MSUs in addition to the regular ambulances. It should be noted that we considered the same MSU scenarios and locations as we did in the companion paper by Amouzad Mahdiraji et al. [7], allowing us to compare and assess the performance of the two models. In MSU1, MSUs are placed near the large municipalities, and in MSU2, MSUs are placed near the sparsely populated areas. Figure 3 (c) and (d) show the MSU locations in the MSU1 and MSU2 scenarios.

We also carried out a sensitivity analysis to further assess the proposed DES model, where we 1) increased the number of available MSUs in each MSU location from one to 10 in the MSU1 and MSU2 scenarios and 2) tripled the number of stroke patients in the SHR. For each scenario, we run the simulation model 10 times with 10 different sets of synthetic stroke patients, taking the averages of the outputs for each of the scenarios.

We generated the driving times using the Openrouteservice toolbox in QGIS, which provides an interface towards open street map (openstreetmap.org). We assumed that a regular ambulance drives 5% faster than a passenger car, and an MSU drives at the same speed as a passenger car. We used Google maps and official documentation provided by the health care authorities in the region to acquire the locations of acute hospitals and ambulance sites. The rest of the assumptions, including the modeled activities, were the same as in the study by Amouzad Mahdiraji et al. [7].

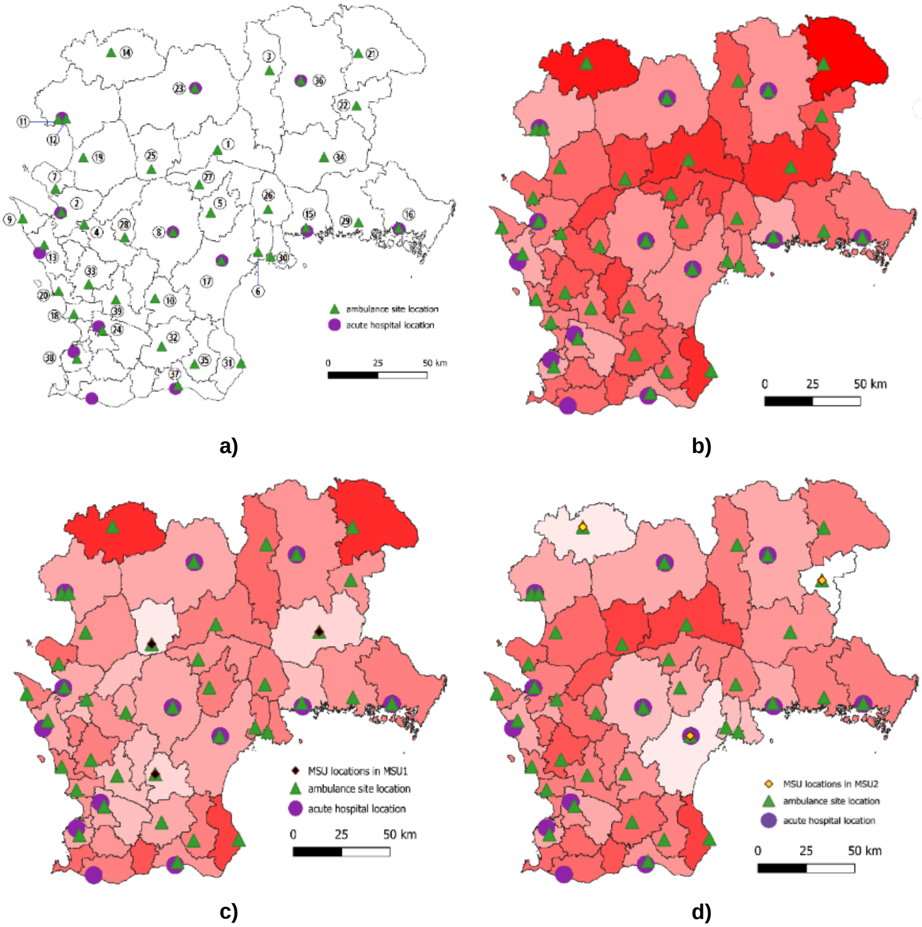


Figure 3. Overview of SHR (a), and the average time to treatment for each of the municipalities of SHR considering the baseline scenario (b), MSU1 (c), and MSU2 (d).

3.1. Synthetic Stroke Population

Due to privacy reasons, there was no access to any individual level patient data for the considered region; we, therefore, used real aggregated stroke data [10] and demographic data [9] in the synthetic population generation model. These data sets, which we used as input to the synthetic stroke generation model, contain the number of inhabitants in each sub-region, as well as the statistical distributions of the time, location, and age of the stroke patients. The model is stochastic, and it generates a new set of patients with varying size, locations, and stroke times, each time it is run.

Our model assumes that the number of stroke incidents for each of the hours of the day is Poisson distributed, and the time between two stroke events is exponentially distributed. The population generation simulation starts by deciding the size of the stroke population and the number of simulation days. To deal with the issue that the expected number of stroke incidents varies over the hours of the day, we first extended each day, in such a way that it consists of a number of time periods of equal length, all having the same number of expected stroke incidents per period. Then, in the extended day, we determined the time of each of the expected stroke incidents. After that, we compressed the extended day again to the initial day length. For each of the generated stroke incidents, we then sampled each patient location and age according to the aggregated statistics.

In our scenario study, we used our synthetic population generation model to create 10 different sets of stroke patients for the SHR, each corresponding to one year. On average, the generated stroke populations contained 3946 patients.

3.2. Results

In Table 1, we present the simulation results of our DES model for each scenario regarding the average time to treatment and the share of the stroke patients who are expected to get treatment within 60, 75, and 90 minutes. The simulation results in Table 1 indicate that after placing MSUs in the SHR, the average time to treatment is expected to decrease for both MSU scenarios compared to the baseline scenario. Furthermore, MSU1 provides faster treatment on average compared to MSU2, which is probably due to that the MSUs in MSU1 are located in more populated regions, having a higher number of stroke incidents than the other regions. In addition, the share of patients who are expected to receive treatment within an hour significantly increased for both MSU scenarios compared to the baseline.

In Figure 3 (b) to (d), we show, for each of the three scenarios, how each MSU in each of the two MSU scenarios is expected to influence the average time to treatment for each municipality in the SHR. The green triangles, purple circles, and black and yellow diamonds illustrate the locations of ambulance stations, hospitals, and MSU locations in MSU1 and MSU2, respectively. It should be emphasized that the lighter the color is in the maps, the shorter the average time to treatment is for the corresponding municipality. It can be further seen that where the considered MSUs are placed, the average time to treatment is reduced in the corresponding municipality as well as in the nearby municipalities.

Table 1. Comparison of the DES model and the macro-level model [7] concerning the average time (in hours) to treatment for the whole SHR and the share of the stroke population in the SHR whose time to treatment is expected to be initiated within 60, 75, and 90 minutes for each scenario. The numbers within the curly brackets show the ambulance site IDs. The simulation runs that belong to the sensitivity analysis are marked with an asterisk (*) in the Model column. ATT: average time to treatment.

Scenario	**MSU sites**	**Model**	**# of MSUs in each site**	**Stroke population**	**ATT**	**Expected time to treatment**		
						<1.0 h	**<1.25 h**	**< 1.5 h**
Baseline	-	Macro model [7]	-	All inhabitants	1.33	3.96%	47.80%	80.90%
		DES	-	3946	1.31	3.85%	46.26%	80.19%
		DES*	-	11453	1.33	3.82%	45.50%	77.98%
MSU1	{11, 27, 36}	Macro model [7]	1	All inhabitants	1.18	13.3%	70.40%	94.0%
		DES	1	3946	1.20	11.48%	66.16%	91.91%
		DES*	10	3946	1.18	13.06%	69.86%	93.80%
		DES*	1	11453	1.25	8.82%	55.64%	86.95%
		DES*	10	11453	1.18	13.25%	68.87%	93.54%
MSU2	{15, 18, 24}	Macro model [7]	1	All inhabitants	1.22	11.70%	58.50%	89.50%
		DES	1	3946	1.23	10.74%	56.08%	88.03%
		DES*	10	3946	1.22	11.66%	57.33%	88.93%
		DES*	1	11453	1.27	8.93%	48.22%	84.22%
		DES*	10	11453	1.23	11.74%	51.95%	87.02%

In Table 1, we also compare the generated results of our DES model with the corresponding results for the macro-level model proposed by Amouzad Mahdiraji et al. [7]. In the current paper, we move forward from a deterministic set-up by proposing a model built on the DES paradigm, which allows to incorporate uncertainty into the simulation set-up. In addition, it allows for a more realistic scenario where the individual patients are simulated over time, and where stroke incidents may occur simultaneously. The proposed DES model considers the availability of ambulances and MSUs, e.g., when there are two simultaneous stroke incidents in a region, the MSU is assigned to one of the stroke incidents; however, for the other stroke incident, the closest regular ambulance will be dispatched instead. In our scenario study, we used the same MSU scenarios as suggested by Amouzad Mahdiraji et al. [7]; however, we applied a synthetic stroke population as input to the DES model.

According to Table 1, the time to treatment is almost equal for the two compared models, i.e., the macro-level and the DES model; however, the DES model shows a slightly longer time to treatment for all scenarios. The reason for the slight differences in the results for the two models probably is that the DES model is able to take into account a limited availability of EVs and the possibility of simultaneous stroke incidents. The more coincident cases there are in the synthetic stroke population, the longer the time to treatment would be.

3.3. Sensitivity Analysis

To further validate our DES model, we conducted a sensitivity analysis, which is presented in Table 1 and in Table 2. In the first part of our sensitivity analysis, we placed 10 MSUs in each site, and in the second part of our sensitivity analysis, we enlarged the synthetic stroke population of SHR three times, i.e., to 11453. Finally, we combined the two parts for the MSU1 and MSU2 scenarios, i.e., we both increased the number of MSUs and the number of patients. When there are 10 MSUs in each site, the average time to treatment decreases for both MSU scenarios compared to the situation where one MSU is placed in each station. When the stroke population is tripled, the average time to treatment increases, and it can be observed that the results for the DES model when having 10 MSUs in each site are equal to the macro-level model results. The reason is that when we increase the number of MSUs to 10, we assure that there is always an MSU available for each stroke case, which is also the case in the macro-level model, which is not able to limit the number of MSUs. If we reduce the number of MSUs to one in each of the ambulance sites, the average time to treatment is longer for the DES model since simultaneous stroke cases sometimes occur in a region, meaning that a regular ambulance has to be dispatched to a patient instead of an MSU. By calculating the difference between the results when having 1 and 10 MSUs in each site, we estimated that the number of patients that could not get an MSU, even though they needed one, was 297 and 124 in MSU1 and MSU2, respectively (see also Table 2). This difference explains why the results of the standard experiment of DES in Table 1 are slightly worse, and most likely more realistic, than for the macro-level model. In addition, when the stroke population is tripled, with the same set-up, the share of MSU dispatches decreases for both MSU scenarios.

In Table 2, we present the number of MSU dispatches and the MSU dispatching ratio, which is defined as the ratio of the total number of MSU dispatches to the total number of EV dispatches for each MSU scenario for the DES model. The simulation results indicate that MSU1 has more MSU dispatches and a higher MSU dispatching ratio than MSU2; hence, the MSUs in MSU1 are expected to help a higher number of stroke patients. Also, by placing 10 MSUs in each EV station, the number of MSU dispatches increases for both MSU scenarios, and again, in this situation, the number is higher for MSU1. In addition, for the standard experiment, the average dispatching distance is 45.38 km and 40.54 km for the MSUs in MSU1 and MSU2, respectively. The probable reason that the average dispatching distance is shorter for MSU2 is that the MSUs in MSU2 are located farther from the densely populated areas; therefore, for most stroke cases, it is more beneficial to dispatch a regular ambulance rather than an MSU.

Table 2. Comparison of the number of MSU dispatches and the share of the total MSU dispatches of each MSU scenario using the DES model in our sensitivity analysis. The numbers within curly brackets show the ambulance site IDs. See Figure 3 (a) for more information.

Scenario	**MSU1**				**MSU2**			
MSU sites	{11, 27, 36}				{15, 18, 24}			
# of MSUs in each site	1	10	1	10	1	10	1	10
Stroke population	3946	3946	11453	11453	3946	3946	11453	11453
Number of MSU dispatches	1651	1948	3385	6185	1017	1141	2217	3299
MSU dispatching ratio (%)	41.84	49.37	29.56	54.00	25.77	28.91	19.36	28.80

4. Conclusions

We have presented a discrete event simulation (DES) model to evaluate MSU-related policies. The model simulates the main actions and decisions involved in stroke transport logistics. From the simulation outputs, we compared the time to treatment, the number of MSU dispatches, and the average dispatching distance of MSUs for the considered scenarios. The simulation results showed that the use of MSUs is expected to lead to a reduced time to treatment in the considered case study, i.e., in southern Sweden, and to help more stroke patients get rapid treatment. Also, with the use of MSUs, the share of patients who are expected to receive treatment within an hour was approximately tripled. Another result was that when MSUs are located in or near the highly populated regions, the number of MSU dispatches is expected to considerably increase, and inhabitants who live quite far from an MSU station are expected to benefit from the MSUs. Finally, by comparing the results of this study with the results of a deterministic model [7], supported by our sensitivity analysis, we conclude that the DES model is able to provide more realistic results than the results that can be obtained using the macro-level modeling approach.

References

[1] J. L. Saver, “Time is brain—quantified,” Stroke, vol. 37, no. 1, pp. 263-266, 2006.

[2] M. Ebinger, B. Winter, M. Wendt, J. E. Weber, C. Waldschmidt, M. Rozanski, A. Kunz, P. Koch, P. A. Kellner, and D. Gierhake, “Effect of the use of ambulance-based thrombolysis on time to thrombolysis in acute ischemic stroke: a randomized clinical trial,” *Jama*, vol. 311, no. 16, pp. 1622-1631, 2014.

[3] H. Zhao, S. Coote, D. Easton, F. Langenberg, M. Stephenson, K. Smith, S. Bernard, D. A. Cadilhac, J. Kim, and C. F. Bladin, “Melbourne Mobile Stroke Unit and Reperfusion Therapy: Greater Clinical Impact of Thrombectomy Than Thrombolysis,” *Stroke*, vol. 51, no. 3, pp. 922-930, 2020.

[4] R. Cerejo, S. John, A. B. Buletko, A. Taqui, A. Itrat, N. Organek, S. M. Cho, L. Sheikhi, K. Uchino, and F. Briggs, “A mobile stroke treatment unit for field triage of patients for intraarterial revascularization therapy,” *Journal of Neuroimaging*, vol. 25, no. 6, pp. 940-945, 2015.

[5] B. M. Bogle, A. W. Asimos, and W. D. Rosamond, “Regional evaluation of the severity-based stroke triage algorithm for emergency medical services using discrete event simulation,” *Stroke*, vol. 48, no. 10, pp. 2827-2835, 2017.

[6] J. Al Fatah, A. a. Alshaban, J. Holmgren, and J. Petersson, “An agent-based simulation model for assessment of prehospital triage policies concerning destination of stroke patients,” in *Proceedings of the 8th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH 2018)*, vol. 141, pp. 405-412, 2018.

[7] S. Amouzad Mahdiraji, O. Dahllöf, F. Hofwimmer, J. Holmgren, R.-C. Mihailescu, and J. Petersson, “Mobile stroke units for acute stroke care in the south of Sweden,” *Cogent Engineering*, vol. 8, no. 1, 2021, doi: https://doi.org/10.1080/23311916.2021.1874084.

[8] The Swedish Stroke Register, “Stroke registrations,” [Online]. Available: https://www.riksstroke.org/sve/forskning-statistikoch-verksamhetsutveckling/statistik/registreringar. [Accessed: Dec. 20, 2019].

[9] Statistics Sweden, “demographic data 2018,” [Online]. Available: https://www.scb.se. [Accessed: Sept. 6, 2018].

[10] Sweden’s Southern Regional Health Care Committee, “stroke data 2018,” [Online]. Available: https://sodrasjukvardsregionen.se. [Accessed: Sept. 6, 2018].

PAPER IV - A FRAMEWORK FOR CONSTRUCTING DISCRETE EVENT SIMULATION MODELS FOR EMERGENCY MEDICAL SERVICE POLICY ANALYSIS

Saeid Amouzad Mahdiraji, Johan Holmgren, Ala'a Alshaban, Radu-Casian Mihailescu, Jesper Petersson, and Jabir Al Fatah

ABSTRACT

Constructing simulation models can be a complex and time-consuming task, in particular if the models are constructed from scratch or if a general-purpose simulation modeling tool is used. In this paper, we propose a model construction framework, which aims to simplify the process of constructing discrete event simulation models for emergency medical service (EMS) policy analysis. The main building blocks used in the framework are a set of general activities that can be used to represent different EMS care chains modeled as flowcharts. The framework allows to build models only by specifying input data, including demographic and statistical data, and providing a care chain of activities and decisions. In a case study, we evaluated the framework by using it to construct a model for the simulation of the EMS activities related to acute stroke. Our evaluation shows that the predefined activities included in the framework are sufficient to build a simulation model for the rather complex case of acute stroke.

PUBLISHED IN

Proceedings of the 12th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH 2022), vol. 210, pp. 133-140, 2022.

1. Introduction

The use of simulation models has become increasingly common for analyzing healthcare systems [1], particularly for the analysis of emergency medical services (EMS) systems, which are emergency services responsible for prehospital stabilization, treatment, and transport of severely injured and sick patients. Simulation models in the healthcare domain are applied to, for example, staff scheduling, allocation of human resources, patient flow management inside emergency rooms, hospital bed utilization, and prehospital healthcare service management [2]. Simulation models are often useful for analyzing complex systems, allowing to evaluate the performance of a system under different configurations and to assess decision policies before being implemented. Discrete event simulation (DES) is a simulation modeling paradigm frequently used to analyze dynamic and complex systems in different areas, such as healthcare. It models the behavior of a system as a sequence of discrete events that occurs over time, where each event takes place at a specific time. The key feature of the DES paradigm is that it only models the points of time where events occur, which can be contrasted to the fixed-increment time simulation approach, where each time step is explicitly simulated.

A policy is typically referred to as a series of activities in the form of guidelines that aims to assist decision-makers in making decisions on a specific problem. Policy assessment in healthcare contributes to identifying the policies that have the potential to help achieving the intended goals and allow the decision-makers to improve the design and implementation of current healthcare policies, leading to an enhanced level of the patients' health and healthcare services. For example, the expected time to treatment is a key factor in certain medical conditions, such as stroke. One of the major objectives of using simulation is to assess healthcare policies that aim to reduce the expected time to treatment for the patients.

A modeling framework typically provides a general structure for modeling a real system and reusable components that can be used and combined in a way that is specified by the user. Modeling frameworks usually provide some types of graphical user interface for constructing models, for example, in the form of an activity diagram. In a framework, a set of activity types are typically integrated to enable activity chains to be modeled as flowcharts or activity diagrams. Activity diagrams or workflow diagrams are behavioral diagrams that represent the workflow of activities from a start point to an endpoint, typically with the support of decisions. Modeling frameworks have been applied in several problem domains, for example, in the transport domain [3-6] and the healthcare domain

[7, 8]. As an example of using a model construction framework in the healthcare domain, Nasir and Kuo [7] contribute a decision support framework to analyze home healthcare (HHC) service delivery decisions of the aging society, allowing to create simultaneous schedules and route plans for both HHC staff and home delivery vehicles. Traoré et al. [8] present a framework for multi-perspective modeling and holistic simulation of healthcare systems. The included perspectives are resource allocation, health diffusion, population dynamics, and individual behavior. Traoré et al. also propose an integrative approach for the interactions between the models of different perspectives and a stratification of the levels of abstraction into multiple perspectives. To the best of our knowledge, no prior study explicitly contributes a modeling framework for the analysis of EMS policies with the focus on prehospital healthcare.

In this paper, we contribute a modeling framework that can be used to build simulation models for EMS policy analysis for different diagnoses. In the rest of the paper, we refer to it as the *model construction framework*. The model construction framework aims to simplify the generation of specific EMS simulation models, by providing an intuitive image of the model to the user, and by providing reusable components that can be used to construct different models. In particular, the framework includes a series of building blocks, which are used to construct care chains modeled as flowcharts. Since decision policies and behaviors of entities are also included in the framework, the framework reduces the need to write code when constructing a new model. To enable the construction of advanced simulation models, we made use of a complex scenario, that is, acute stroke, to identify the components to include in the framework.

The rest of the paper is organized as follows. In Section 2, we describe the proposed model construction framework for constructing EMS simulation models, followed by the scenario study in Section 3, where we applied the proposed framework to build a simulation model for the stroke care chain. Eventually, we conclude the paper in Section 4.

2. The Model Construction Framework

In this section, we present our model construction framework, which can be used to construct EMS simulation models for different medical conditions. The framework is a micro-level approach, meaning that it can be used to build models where each patient is simulated individually.

In general, it is complicated to build one simulation model that captures all aspects of the EMS domain, and a model therefore typically needs to be created

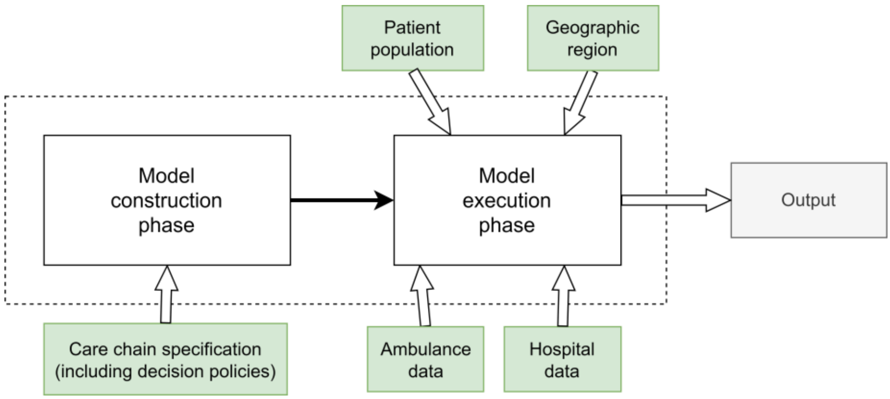


Figure 1. Overview of the proposed model construction framework, which is an integrated approach where the user provides input data that is used to 1) construct the model and 2) run the model.

for the purpose of analyzing only one or a few aspects; for example, the analysis of patient admission process policies [9], dynamic ambulance dispatching policies [10], stroke-related policies [11, 12], or prediction of medical workforce supply [13]. However, different EMS models share many aspects; hence, it is often a good idea to build models using a reusable framework rather than building them from scratch. The framework enables to create a model only by specifying the care chain and providing input data. Hence, the framework simplifies the generation of specific models and reduces the complexity of the model construction activities.

A schematic overview of our model construction framework is presented in Figure 1. The framework is an integrated approach, in the sense that it contains both a model construction phase and a model execution phase. The user provides different types of input, which are used to configure the framework and to provide the necessary specifications for the modeled entities. In other words, the model construction framework utilizes input, such as statistical data, demographic data, and a care chain specification, including activities and decisions, to build EMS simulation models, which are then automatically run using the input data. Examples of output generated by the framework are the times to diagnosis and treatment for each patient, and the total service time of the ambulances.

We studied typical EMS care chains to include the most common activities in our framework. The identified activities include, for example, *choose destination hospital*, *transport patient*, and *diagnose patient*, and each model activity is connected to one of the modeled actors, that is, *emergency call center*, *ambulance*, and *hospital*. See Table 1 for a complete list of the general activities provided by the framework. Using the provided activities, we argue that the framework can be used to construct most chains of main activities and decisions from when the emergency call center receives an emergency call until the patient receives treatment.

We regard an activity as a key task that should be executed as part of completing a process. Each activity has its corresponding actor and input, functionality, and output. Each activity in a chain of activities typically also requires some specific prerequisites to be triggered, that is, the previous activity must be completed, and a possible condition for starting that activity might need to be fulfilled. There are two categories of activities: decision-making activities, for example, for the ambulance or hospital selection, and pure task activities, for example, *transport* and *diagnosis*.

The framework makes use of different types of input, that is, geographic region, patient population, ambulance data, hospital data, and a care chain specification. In the following, we describe each type of input in detail.

Table 1. List of general care chain activities that are included in the framework.

Activity	Actor	Description
Emergency call	Emergency call center	The emergency call center receives a call concerning an emergency incident.
Preassessment	Emergency call center	The emergency call center performs preassessment over the phone. Based on the description of the patient symptoms, an ambulance priority and a set of possible diagnoses are determined.
Select ambulance	Emergency call center	If an ambulance is required, the emergency call center decides which ambulance should be dispatched to provide service to a patient.
Transport	Ambulance	In general, an ambulance may travel using three different paths in the EMS care chain: Travel from the ambulance current location to the patient scene. Travel from the patient scene to a hospital. Travel from a hospital to an ambulance site.
Patient scene assessment	Ambulance	The ambulance paramedics examine the patient at the patient scene, determine a set of possible diagnoses, or/and exclude some diagnoses.
Load patient	Ambulance	The patient is loaded into the ambulance.
Choose destination hospital	Ambulance	Based on the patient symptoms, the paramedics' diagnosis, and predefined policies, the paramedics choose to which hospital the patient should be transported.
Unload patient	Ambulance	The patient is unloaded from the ambulance.
Diagnosis	Hospital	The patient is examined at the hospital.
Treatment	Hospital/ Ambulance	The treatment is initiated in the hospital or in the ambulance.

Geographic region: A model for the analysis of EMS care chains needs to explicitly model a geographic region, which is populated by hospitals, patients, and ambulances. The framework assumes that the geographic region is divided into smaller subregions, where each subregion is represented by its id and its centroid coordinates (longitude, latitude). It is also assumed that all inhabitants residing in a subregion are represented by its centroid, where the centroid is the core of all activities of the subregion it represents.

Patient population: The simulation is driven forward by patients appearing at different times and locations in a geographic region. The patient population can be real or synthetic, where each patient is specified using attributes such as age, sex, incident time (day and hour), location (coordinates), and symptoms.

Hospital and ambulance data: Hospital data includes hospital locations (coordinates). Ambulance data contains ambulance site locations (coordinates) and expected traveling times between 1) ambulance sites and patient locations, 2) patient locations and hospital locations, and 3) hospital locations and ambulance sites.

Care chain specification: The framework uses building blocks to construct care chains of EMS activities modeled as flowcharts. The building blocks are general reusable activities representing both decision-making and non-decision-making activities. The flowchart approach provides an intuitive way for the user to construct complex care chains, where the care chain flowchart used in our approach consists of boxes, arrows, an ellipse, and conditions, where two consecutive blocks are connected to each other using an arrow. As an example, in Figure 2, we show the chain of EMS activities for a suspected stroke patient, where each block specifies an activity and its corresponding actor. An arrow represents the order of activities, and the ellipse represents a trigger to start the care chain, which in the considered scenario is the emergency call. A condition specifies a condition that must be met to go from one activity to another, here shown by a note connected to the arrow. For example, the note *Condition: patient.worst_case_diagnosis_code* between the *preassessment* and *select ambulance* activities in Figure 2, implying that if an ambulance assignment is decided, the *select ambulance* activity will begin; otherwise, the care chain for the corresponding patient will be stopped when the *preassessment* activity is executed. At runtime, the framework constructs a specific simulation model, where the provided care chain specification is applied for each of the simulated patients.

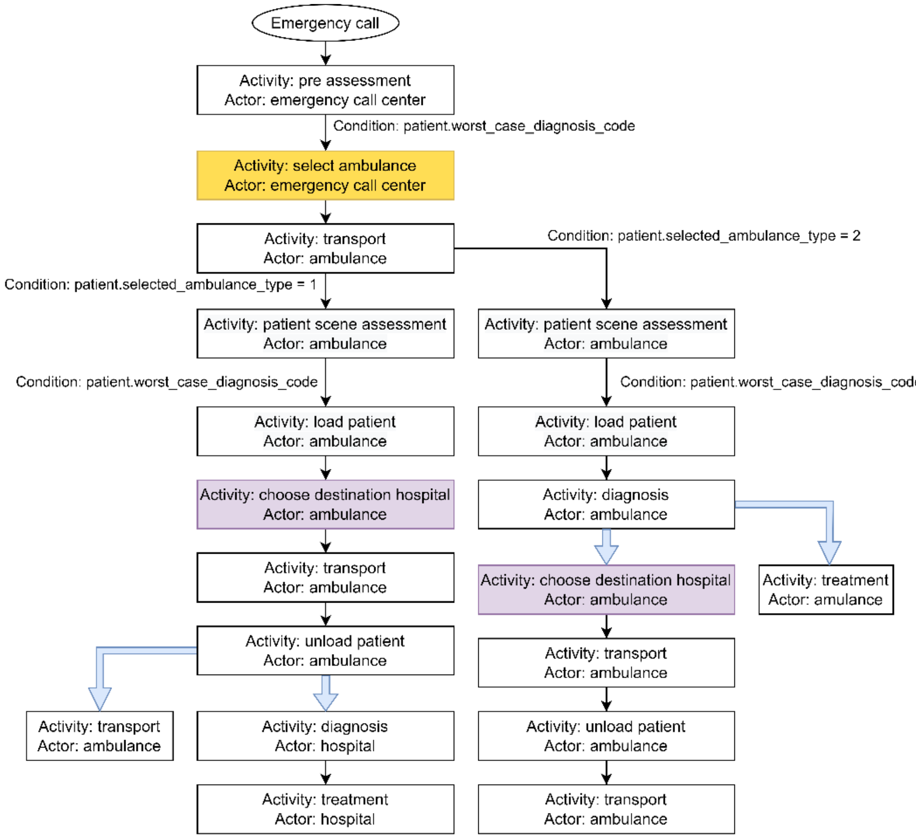


Figure 2. Care chain of the EMS activities for a suspected stroke patient, where the colored boxes demonstrate the decision-making activities and the blue arrows represent the parallel activities.

We have also included parallel activities that typically occur in the EMS care chains, shown by the blue arrows in Figure 2. For example, after unloading the patient from the ambulance (*unload patient* activity), the patient will be prepared for the diagnosis at the hospital, and the ambulance will simultaneously return back to its station.

The major decisions currently modeled in the framework, and which are included in the care chain example in Figure 2 are 1) which ambulance should be dispatched towards the patient? and 2) to which hospital should the patient be transported? In a typical situation, the above decisions can be made by assigning the closest available ambulance for the operation and transporting the patient to the nearest hospital from their current location. However, there may be exceptions, for example, when a special type of ambulance is required based on the preassessment over the telephone by the emergency call center. In such a case, a special ambulance needs to be dispatched instead. Or, after the patient scene assessment by paramedics, the patient needs to be transported to a hospital that can provide specific treatment. In this case, the patient will be directly driven to the special clinic. Based on the above reasoning, we included a set of different policies for selecting ambulance and hospital, which is discussed in more detail below.

The ambulance selection policies aim to choose an emergency vehicle (EV) and its type by minimizing 1) the expected time to patient, 2) the time to diagnosis, or 3) the time to treatment, represented using the following policies:

- **Time to patient policy**: An EV is chosen to minimize the expected time from symptoms onset until the EV arrives at the patient location.
- **Time to diagnosis policy**: An EV is chosen to minimize the expected time from when an incident happens until the diagnosis is initiated at the hospital or patient location.
- **Time to treatment policy**: An EV is chosen to minimize the expected time from when an incident happens until the patient receives treatment at the hospital or at the patient location.

There are typically different types of EVs available, which makes it reasonable to also include the type of EV in the ambulance selection decision. Each ambulance selection policy enables choosing the EV that can provide the fastest service for the patient. By conducting a set of simulation runs, it is then possible to compare different decision policies to observe from which policy the patient is expected to benefit most. It should be highlighted that the time to diagnosis and the time to treatment policies are made using assumptions based on the emergency call center preassessment.

To decide where to transport the patient (to which hospital), the following policies are included in the framework:

- **Closest hospital policy**: The patient is transported to the closest hospital from their location.
- **Direct to special clinic policy**: The patient is transported directly to the closest special clinic that can provide specific treatment based on the outcome of the patient scene assessment.

It should be emphasized that the abovementioned policies are the policies currently coded into our framework to choose an ambulance and hospital for a patient. If other policies are needed, they can be easily added to the framework.

As an example, in Figure 2, the decision for selecting an EV and its type is made in the *select ambulance* activity colored by orange, using the predefined ambulance selection policy. The decision for choosing the destination hospital is made in the *choose destination hospital* activity, colored by purple, using the predefined hospital selection policy.

We established our framework on the principles of the DES paradigm, meaning that the framework, based on the specified entities, constructs a DES model for the analysis of the EMS care chain. Each activity in the care chain is represented using two events, that is, a start event and an end event, enabling to simulate activities that are non-instantaneous. Upcoming events are sorted in a queue based on the time of occurrence, meaning that the next event is always the first event in the queue. Once an event occurs, the simulation clock is moved forward to the time of that event, where it represents either the start or the end of an activity. As shown in Figure 2, the care chain of activities for each patient is initiated by creating an *emergency call* activity. At the end of an activity, the start event for the next activity is typically created and added to the event queue. For example, when the *emergency call* activity is handled, the model creates the start event for the *preassessment* activity and adds the new event in the right place in the queue. When an event has been processed, it is removed from the queue. The simulation will terminate once the event queue is empty or when the simulation end-time has been reached.

3. Scenario Study

We evaluated our framework by constructing a model for the simulation of EMS care chain activities for stroke patients. In this section, we initially provide a brief introduction to stroke, and we then present the application of our framework to construct a simulation model for the stroke care chain.

Stroke is one of the deadliest diseases and a major cause of disability around the world. Thrombolysis is the standard treatment for ischemic stroke, which is the most common type of stroke. It is generally agreed that the time to treatment is the most crucial factor for the possibility of stroke patients to recover. However, logistical issues typically make it difficult to provide treatment fast enough, as stroke patients traditionally need to be transported to an acute hospital for diagnosis and treatment However, it has been shown that it is possible to reduce the time to treatment for stroke patients by utilizing so-called Mobile Stroke Units (MSUs), which are special ambulances equipped with a CT scanner, allowing to both diagnose stroke patients and provide thrombolysis inside the MSU [14]. As mentioned earlier, the use of simulation can help us choose decision policies that lead to reduced expected time to treatment for stroke patients. For the case of MSUs, an important decision policy concerns where to locate the MSUs to provide the best possible service for the inhabitants in a geographic region.

We chose acute stroke for our scenario study since we consider the stroke care chain to be rather complex in comparison to the EMS care chain for many other diagnoses. An important reason for this is that the use of MSUs introduces multiple decisions that should be taken as there might be different types of EVs available to provide service to the patient. Furthermore, decisions often need to be taken based on incomplete information, for example, since the patient cannot be diagnosed until after conducting a CT scan on the brain of the patient, it is challenging to decide where to transport the stroke patient, that is, the closest hospital or special clinic.

In our scenario study, we used both regular ambulances and MSUs in the EMS care chain for stroke patients in need of thrombolysis. The considered care chain, which is illustrated in Figure 2, is triggered when the emergency call center receives a call concerning an emergency incident, and the subsequent activities will be activated during the simulation. Each activity is triggered once the preceding activity is accomplished and in case the possible condition on the incoming arrow is met. Otherwise, the care chain terminates after the current activity. For example, the conditioned arrow with the note *Condition: patient.worst_case_diagnosis_code* between the *patient scene assessment* and *load patient* activities indicates that the paramedics determine whether the patient should be transported to the hospital for treatment based on the worst-case code given based on the patient symptoms. If it is decided to transport the patient, the patient is loaded into the ambulance (*load patient* activity), and the following care chain activities will be executed. Otherwise, the care chain for the corresponding patient will be stopped after the *scene assessment* activity has been executed. Then, the patient is left at their current location, and the ambulance returns back to its station.

When an MSU is dispatched for a potential stroke case, the diagnosis and treatment (thrombolysis) can be performed at the patient scene. In this case, the time to diagnosis/treatment is the expected time from when a stroke happens until the patient receives diagnosis/treatment inside the MSU. However, when a regular ambulance is dispatched, the patient should be transported to a hospital for diagnosis and treatment. After unloading the patient at the hospital, the EV returns to its station, and once it arrives there, it becomes available for the next operation.

In the stroke care chain, there are two situations when activities happen in parallel, represented by the blue arrows in Figure 2. When a regular ambulance is assigned for an operation, there is one pair of parallel activities, which is described in Section 2. There is also a pair of parallel activities when an MSU is assigned for an operation, see the blue arrows in the right side of Figure 2, showing that after performing the diagnosis inside the MSU, the treatment is initiated at the same time as the paramedics choose the destination hospital, followed by the MSU starting to drive towards the hospital.

In our scenario study, we constructed a model for the EMS care chain of stroke patients in Sweden's Sothern Healthcare Region (SHR) and conducted a set of simulation runs using the constructed model. In the simulation runs, we considered all of the decision policies described in Section 2. The constructed simulation model takes as input geographic region data, patient population, and hospital and ambulance data. We also provided data related to the acute hospitals and special clinics that can perform thrombectomy treatment. We used the same input data, MSU locations, and assumptions as we did in a previous study [11]. The geographic region is the SHR, where all patients, ambulance sites, and hospitals are located, and the geographic region is divided into 1×1 km squares. The stroke patient data is a synthetic population of patients distributed over the SHR, and it was generated using a Poisson distribution (see Amouzad Mahdiraji et al. [11] for details). We assumed that all modeled patients are suspected stroke patients, which need to be transported to a hospital. The scenario study included three different scenarios: a baseline scenario (including only regular ambulances) and two MSU scenarios (MSU1 and MSU2), each including three MSUs. It was assumed that the MSUs in MSU1 are located in Hörby, Markaryd, and Tingsryd, and the MSUs in MSU2 are located in Lessebo, Hyltebruk, and Kristianstad (see Amouzad Mahdiraji et al. [15] for details).

In Table 2, we present the simulation results for the constructed model, in the form of expected times to treatment (in hours), for each scenario and each of the ambulance and hospital selection decision policies described in Section 2. As expected, the simulation results show that the use of MSUs, compared to the baseline scenario, enables to provide better service for stroke patients in the SHR by reducing the expected time to treatment for both of the considered MSU allocations. In particular, when the decision for the choice of an EV is made using each of the *time to diagnosis* and *time to treatment* policies. The improvement is

Table 2. Comparison of the average expected time to treatment (in hours) obtained from simulation model of stroke care chain for the whole SHR for each scenario and decision policy. The numbers within the curly brackets show the ambulance site IDs. See our prior paper [15] for details. TP: time to patient, TD: time to diagnosis, TT: time to treatment.

Scenario	MSU sites	Direct to special clinic			Closest hospital		
		TP	TD	TT	TP	TD	TT
Baseline	-	2.020 h	2.020 h	2.020 h	1.347 h	1.347 h	1.347 h
MSU1	{11, 27, 36}	2.017 h	1.571 h	1.584 h	1.346 h	1.279 h	1.274 h
MSU2	{15, 18, 24}	2.014 h	1.602 h	1.618 h	1.346 h	1.291 h	1.287 h

more considerable when the hospital selection policy is *direct to special clinic*; especially, the expected time to treatment is expected to be reduced from 2.020h for the baseline scenario to 1.571h and 1.602h for the MSU1 and MSU2 scenarios, respectively, considering the *time to diagnosis* policy.

For the *closest hospital* policy, the share of patients who are expected to receive treatment within an hour is approximately doubled and tripled for the MSU1 and MSU2 scenarios, respectively. In addition, the share of the patients who are likely to receive a diagnosis within an hour is expected to increase from 3.65% for the baseline to 90.91% and 88.05% for MSU1 and MSU2, respectively. Considering our assumption that a regular ambulance drives faster than an MSU, the results in Table 2 suggest that when the decision to select an EV depends on the *time to patient* policy, the use of a regular ambulance is a more effective choice. The results presented in Table 2 are similar to the results presented in our prior paper [11], where we employed a DES model to evaluate MSU-related policies in the stroke care chain, demonstrating that the constructed simulation model for the stroke care chain functions as intended.

4. Conclusions

In this article, we proposed a model construction framework that can be used to simplify the construction of EMS simulation models. The framework is based on the idea of modeling EMS care chains using standard flowcharts consisting of connected activities and decisions, and it includes a set of general activities that can be used to represent most EMS care chains. As the framework includes activities and policies modeled on a general level, the framework can construct models only by specifying input, such as geographic region, patient population, hospital data, ambulance data, and the care chain. We evaluated our framework by applying it to construct a simulation model for the analysis of EMS activities for the stroke care chain, which we argue is rather complex in comparison to the typical EMS care chain. The reason we argue that the stroke care chain is more complex is that including an MSU in the prehospital stroke care chain creates multiple decisions that need to be taken with incomplete information; for example, decisions related to EV selection, diagnosing the type of stroke, and hospital selection. After applying the required data to the generated simulation model for stroke patients, the simulation output analysis showed that the included activities in the framework were sufficient to build EMS simulation models, representing complex diagnoses such as stroke.

References

[1] M. R. Davahli, W. Karwowski, and R. Taiar, "A system dynamics simulation applied to healthcare: A systematic review," *International Journal of Environmental Research and Public Health*, vol. 17, no. 16, 2020.

[2] S. Almagooshi, "Simulation modelling in healthcare: Challenges and trends," . *Procedia Manufacturing*, vol. 3, pp. 301-307, 2015.

[3] Q. Deng, "A general simulation framework for modeling and analysis of heavy-duty vehicle platooning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 11, pp. 3252-3262. 2016.

[4] J. Holmgren, P. Davidsson, J. A. Persson, and L. Ramstedt, "TAPAS: A multi-agent-based model for simulation of transport chains," *Simulation Modelling Practice and Theory*, vol. 23, pp. 1-18, 2012.

[5] E. López-Neri, A. Ramírez-Treviño, and E. López-Mellado, "A modeling framework for urban traffic systems microscopic simulation," *Simulation Modelling Practice and Theory*, vol. 18, no. 8, pp. 1145-1161, 2010.

[6] K. F. Abdelghany, H. S. Mahmassani, and A. F. Abdelghany, "A modeling framework for bus rapid transit operations evaluation and service planning," *Transportation Planning and Technology*, vol. 30, no. 6, pp. 571-591, 2007.

[7] J. A. Nasir and Y. -H. Kuo, "A decision support framework for home health care transportation with simultaneous multi-vehicle routing and staff scheduling synchronization," *Decision Support Systems*, vol. 138, 2020.

[8] M. K. Traoré, G. Zacharewicz, R. Duboz, and B. Zeigler, "Modeling and simulation framework for value-based healthcare systems," *Simulation*, vol. 95, no. 6, pp. 481-497, 2019.

[9] H. Kang, H. B. Nembhard, C. Rafferty, and C. J. DeFlitch, "Patient flow in the emergency department: a classification and analysis of admission process policies," *Annals of emergency medicine*, vol. 64, no. 4, pp. 335-342, 2014.

[10] C. J. Jagtenberg, S. Bhulai, and R. D. van der Mei, "Dynamic ambulance dispatching: is the closest-idle policy always optimal?," *Health care management science*, vol. 20, no. 4, pp. 517-531, 2017.

[11] S. Amouzad Mahdiraji, J. Holmgren, R. -C. Mihailescu, and J. Petersson, "A Micro-Level Simulation Model for Analyzing the Use of MSUs in Southern Sweden," in *Proceedings of the 11th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH 2021)*, vol. 198, pp. 132-139, 2022.

[12] J. Al Fatah, A. A. Alshaban, J. Holmgren, and J. Petersson, “An agent-based simulation model for assessment of prehospital triage policies concerning destination of stroke patients,” in *Proceedings of the 8th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH 2018)*, vol. 141, pp. 405-412, 2018.

[13] M. A. Lopes, Á. S. Almeida, and B. Almada-Lobo, “Forecasting the medical workforce: a stochastic agent-based simulation approach,” *Health care management science*, vol. 21, no. 1, pp. 52-75, 2018.

[14] M. Ebinger, B. Winter, M. Wendt, J. E. Weber, C. Waldschmidt, M. Rozanski, A. Kunz, P. Koch, P. A. Kellner, and D. Gierhake, “Effect of the use of ambulance-based thrombolysis on time to thrombolysis in acute ischemic stroke: a randomized clinical trial,” *Jama*, vol. 311, no. 16, pp. 1622-1631, 2014.

[15] S. Amouzad Mahdiraji, O. Dahllöf, F. Hofwimmer, J. Holmgren, R. -C. Mihailescu, and J. Petersson, “Mobile stroke units for acute stroke care in the south of Sweden,” *Cogent Engineering*, vol. 8, no. 1, 2021, doi: https://doi.org/10.1080/23311916.2021.1874084.

PAPER V - AN OPTIMIZATION MODEL FOR THE PLACEMENT OF MOBILE STROKE UNITS

Saeid Amouzad Mahdiraji, Muhammad Adil Abid, Johan Holmgren, Radu-Casian Mihailescu, Fabian Lorig, and Jesper Petersson

ABSTRACT

Mobile Stroke Units (MSUs) are specialized ambulances that can diagnose and treat stroke patients; hence, reducing the time to treatment for stroke patients. Optimal placement of MSUs in a geographic region enables to maximize access to treatment for stroke patients. We contribute a mathematical model to optimally place MSUs in a geographic region. The objective function of the model takes the tradeoff perspective, balancing between the efficiency and equity perspectives for the MSU placement. Solving the optimization problem enables to optimize the placement of MSUs for the chosen tradeoff between the efficiency and equity perspectives. We applied the model, to the Blekinge and Kronoberg counties of Sweden to illustrate the applicability of our model. The experimental findings show both the correctness of the suggested model and the benefits of placing MSUs in the considered regions.

PUBLISHED IN

Proceedings of the International Conference on Advanced Research in Technologies, Information, Innovation and Sustainability (ARTIIS 2023), pp. 297-310, 2023.

1. Introduction

A stroke refers to when a blood clot or a bleeding interrupts the blood circulation inside the brain, and stroke is a main global reason for death and permanent disability [1]. There are three main stroke types, each requiring specific treatment. To assure providing the correct stroke treatment, a computed tomography (CT) scan is required to identify which type of stroke the patient is suffering. Ischemic strokes are most common, and they occur when blood clot(s) impede blood circulation in the brain. Treatment for ischemic stroke patients typically involves thrombolysis and, in specific cases, thrombectomy.

Early treatment is known to be crucial for the successful recovery of stroke patients [2]. However, it is often difficult to treat stroke patients immediately since the patient typically cannot be diagnosed and treated until the ambulance delivers him/her to an acute hospital.

One effective way to decrease the time to treatment for stroke patients involves the utilization of Mobile Stroke Units (MSUs). An MSU is a specialized ambulance equipped with advanced medical equipment, including a CT scanner, that enables the ambulance personnel to diagnose and administer thrombolysis on stroke patients while the patient is still in the MSU. As a result, MSUs have the potential to reduce the time to treatment by eliminating the time required for the transportation and diagnosis of the patient at the hospital. However, due to the high operational expenses associated with MSUs, only a limited number of MSUs can be placed in a geographic region. Therefore, when introducing MSUs, it becomes essential to strategically place them in order to provide a timely service for residents living in a region.

There are a number of studies focusing on identifying the optimal locations for placing MSU(s) within a region to enhance stroke care. These studies mainly explore two perspectives regarding where to place MSUs: efficiency and equity. The term efficiency refers to placing MSUs so that they provide access to treatment in the shortest possible time for most patients in a region, for example, in urban areas [3,4]. Equity emphasizes the placement of MSUs in a manner that ensures equal access to healthcare services regardless of the geographic location of the patients, for example, in rural areas [5]. Phan et al. [3] introduce a data-driven approach that utilizes the Google application programming interface to determine the best possible placement for an MSU within the Sydney area. Rhudy Jr. et al. [4] use geospatial analysis to optimize service delivery for stroke patients in Memphis by studying the distribution of an MSU throughout the city. Dahllöf et al. [6] propose an expected value optimization approach to determine the best placement for an MSU in Sweden's Skåne county, aiming to assess the potential advantages of placing an MSU for urban and rural residents respectively. Amouzad Mahdiraji et al. [7] utilize an exhaustive search approach to optimally place MSUs in southern Sweden with the aim of balancing the efficiency and equity perspectives for the placement of MSUs.

The aim of the current study is to introduce a mathematical optimization model in the form of a mixed integer linear programming (MILP) model to identify the best locations of MSUs within a geographic region. Mathematical optimization has been demonstrated to be an effective technique to solve complex problems in a wide range of domains, such as emergency medical services (EMS). Due to the computational complexity of emergency vehicle placement problems, it is vital to build efficient mathematical models to represent the key characteristics of the MSU placement problem. However, no existing research directly addresses the mathematical formulation of the MSU placement problem. The objective function of the presented MILP model expresses a tradeoff between the efficiency and equity perspectives, aiming to provide maximum population coverage as well as equal service for the inhabitants of a region; however, considering the chosen tradeoff between the two perspectives. A scenario study is conducted in two counties of southern Sweden to show the correctness and advantages of our proposed model, where we solve the model to identify the optimal placements for different numbers of MSUs.

The subsequent sections of the paper are outlined as follows. We review the related work in Section 2. In Section 3, we present the MSU placement problem with a tradeoff between the efficiency and equity perspectives. In Section 4, we present our optimization model for the described problem. The scenario study is presented in Section 5, which is followed by an analysis of experimental results and a discussion. Eventually, we conclude the paper in Section 6.

2. Related Work

Previous studies in EMS use MILP for problems related to ambulance routing and placement, ambulance fleet allocation, crew scheduling, and resource allocation, ultimately leading to better patient outcomes and resource utilization [8]. As an example, Tavakoli et al. [9] propose a mathematical model for the strategic placement of ambulances, aiming to improve the response time of EMS in Fayetteville, North Carolina. Røislien et al. [10] use mathematical modeling to explore the optimal locations for air ambulance sites in Norway. Their approach utilizes high-resolution population data to estimate the number of required sites to provide service within 30 and 45 minutes for different shares of the population. Leknes et al. [11] present a MILP model to address the strategic and tactical problems of placing ambulance sites in heterogeneous regions. The authors examined the model in an urban-rural area in Norway. Akdoğan et al. [12] utilize queuing theory and a MILP model to locate emergency vehicles on fully connected networks. The MILP model aims to reduce the average response time of EMS according to an approximate queuing model.

In another study, Tlili et al. [13] propose a mathematical model to improve EMS transportation during disaster situations. The authors use a genetic algorithm for the ambulance routing problem to reduce time-sensitive treatment delays during urgent situations involving congested traffic compounds. Acuna et al. [14] contribute an ambulance placement optimization model to decrease patients' waiting times, time to treatment, and emergency department overcrowding in a county in Florida. The model considers disparities and fairness in placing ambulance services to emergency departments. Wan S. et al. [15] use a 0-1 MILP model to represent the location of distribution centers in massive emergencies, applied in a case study of earthquake response logistics in Chengdu, China. The proposed bi-objective model considers both the total transportation cost and the coverage level of emergency supplies.

While numerous research studies focus on the mathematical modeling of ambulance location problems, no previous study explicitly contributes to the mathematical formulation of the MSU placement problem. To address this gap, in this paper, we present a MILP model to represent the MSU placement problem.

3. MSU Placement Problem

As mentioned earlier, when placing MSU(s) in a geographic region, we need to take the impacts of the MSU locations into account to assure that the inhabitants of different parts of a region receive maximum benefit. In this section, we describe the MSU placement problem and how MSUs can be placed in a region considering the tradeoff between the efficiency and equity perspectives.

In our companion study [7], we demonstrate how different placements of MSUs would impact individuals living in different parts of a region. In particular, we propose an objective function that could be used in an optimization model to tradeoff between the efficiency and equity perspectives, and hence, allows placing MSUs so that most people living in a region are expected to receive more equitable service and shorter time to treatment. In addition, we employ the concept of the expected time to treatment to capture the value of the corresponding measure for each perspective. It should be noted that the expected time to treatment for a stroke patient denotes the expected time until the patient gets treatment either at a hospital or inside an MSU. In a previous study [16], we present how to calculate the expected time to treatment for patients in different subregions of a geographic region, considering that both a regular ambulance and an MSU can be dispatched.

The efficiency perspective refers to placing MSUs in a region to ensure a higher proportion of the population is expected to receive treatment at an earlier time. Using this perspective, the MSUs are placed close to highly populated regions, that is, in or near the urban areas. The efficiency perspective can be measured by the weighted average time to treatment (WATT). The expected time to treatment for individuals located in each subregion of a larger region is multiplied by the share of stroke cases expected to take place in the corresponding subregion; the sum of these values yields the WATT. We can use the WATT as an objective function in an optimization model for the MSU placement problem that considers the efficiency perspective.

The equity perspective refers to placing MSUs where the people who live far from the medical centers (for example, hospitals) benefit most, that is, people living in or close to rural areas. The range measure can be utilized to model the equity perspective, aiming to minimize the time difference between the expected times to treatment for patients who are located in different subregions of the studied region. The focus of an optimization problem corresponding to the equity perspective is to identify the MSU placements that minimize range.

In our companion study [7], we also introduce a tradeoff function that is established based on the WATT and range. It is shown that the tradeoff function enables to balance between the efficiency and equity perspectives to optimally place MSUs. In an optimization problem for placing MSUs in a region, the tradeoff perspective aims to find the locations of MSUs that minimize the tradeoff function.

It should be highlighted that in the formulated optimization problem, we only consider, for each perspective, the placement of MSUs in the existing ambulance sites in a geographic region.

4. Optimization Model

We here present our MILP model, which represents the key characteristics of the MSU placement problem. Our optimization model aims to minimize the tradeoff function that enables to identify the optimal locations for MSU(s) that can provide highly equitable service and reduced time to treatment for residents within a region.

We let $I = \{1, \dots, m\}$ denote the index set over ambulances sites, where m is the total number of ambulance sites, and N is the number of MSUs to place. The aim is to place a fixed number of MSUs at the existing ambulance sites in a geographic region. It is assumed that there is always at least one regular ambulance available at each ambulance site. We also assume that there is always an ambulance or an MSU available for dispatch when it is required, and that the placed MSU(s) have no limitation concerning driving distance, and that they can provide service throughout the whole region.

We further assume that the studied region is divided into a non-overlapping set of subregions, denoted by $R = \{1, \dots, n\}$, where n is the total number of subregions. We also assume that all inhabitants located in subregion $r \in R$ are in the same location, for example, in the centroid of r.

We let t_r^{RA} be the shortest time to treatment using a regular ambulance located in any ambulance site $i \in I$ for subregion $r \in R$, t_{ir}^{MSU} be the expected time to treatment for a patient located in subregion $r \in R$ using an MSU located in site $i \in I$, and Q_r be the share of stroke incidents within the studies region that is expected to take place in subregion $r \in R$ ($\sum_{r \in R} Q_r = 1$). Please note that the t_r^{RA}:s ($r \in R$), t_{ir}^{MSU}:s ($i \in I$, $r \in R$), and Q_r:s ($r \in R$) are input parameters, and hence can be calculated beforehand.

In order to formulate the MILP model, we need the following decision variables:

- $x_i \in \{0,1\}, (i \in I)$ is a binary decision variable such that:

$$x_i = \begin{cases} 1 & \text{if there is an MSU in site } i \in I, \\ 0 & \text{Otherwise.} \end{cases} \tag{1}$$

- y_{ir}^{MSU} is the expected time to treatment for a patient in subregion $r \in R$ using an MSU in site $i \in I$. This variable is assigned a large value, M, if there is no MSU placed in site $i \in I$.
- y_r^{MSU} is the shortest expected time to treatment for a patient in subregion $r \in R$ using any of the placed MSUs.
- y_r is the shortest expected time to treatment for a patient in subregion $r \in R$ using either an MSU or a regular ambulance.
- u^{max} is the longest expected time to treatment for any subregion $r \in R$.
- u^{min} is the shortest expected time to treatment for any subregion $r \in R$.

The tradeoff function z, presented in Equation (2), is the objective function for our MILP model. The objective function has two components: the first one is the WATT as a measure for the efficiency perspective, and the second one is the range (time difference between subregions with the shortest and longest expected time to treatments) as a measure for the equity perspective.

$$min\ \ z = \sum_{r=1}^{R} (1-w) y_r Q_r + w(u^{max} - u^{min}), \tag{2}$$

In Equation (2), $w \in [0,1]$ is the weight employed to control the effects of the efficiency and equity perspectives. For example, we here assume $w = 0.5$ to let each of the terms have an equal impact on the tradeoff function.

The optimal solution of our model is subject to the following constraints:

$$y_{ir}^{MSU} = \ x_i t_{ir}^{MSU} + M(1 - x_i),\ i \in I,\ r \in R, \tag{3}$$

$$y_r^{MSU} = \ \min_{i \in I} \left\{ y_{ir}^{MSU} \right\}, \qquad i \in I, r \in R, \tag{4}$$

$$y_r = min\ \{ y_r^{MSU}, t_r^{RA} \},\ r \in R, \tag{5}$$

$$u^{max} \geq y_r,\ r \in R, \tag{6}$$

$$u^{min} \leq y_r, \qquad r \in R, \tag{7}$$

$$\sum\nolimits_{i \in I} x_i = N. \tag{8}$$

We use constraint sets (3)-(5) to obtain the values of the y_r, which is the shortest expected time to treatment for any subregion $r \in R$ using either an MSU or a regular ambulance. The constraint in Equation (3) assigns t_{ir}^{MSU} to y_{ir}^{MSU} if there is an MSU available in site i for a patient in subregion r. However, if no MSU is located in site i, it instead assigns a large value M to the y_{ir}^{MSU}. M, which is a parameter in our optimization model, is a sufficiently large constant value. For example, M can be set to any value larger than the longest expected time to treatment for any subregion r and any ambulance site i, that is, $M > \max_{r \in R}\ t_r^{RA}$.

The constraint in Equation (4) takes the minimum over the expected times to treatment for the possible MSU locations. The minimum operation is used to assign the shortest expected time to treatment using an MSU for a patient in subregion r. In the optimization model, the constraint $y_r^{MSU} = \ \min_{i \in I} \{ y_{ir}^{MSU} \} =$

$\min\left\{y_{ir}^{MSU},\ldots,y_{mr}^{MSU}\right\}$ is modeled as an ordered sequence of $(|I|-1)$ minimum operations, each having two components. For this purpose, we introduce a set of positive help variables $p_{ir}^{MSU}, i \in \{1,\ldots,|I|-1\}, r \in R$, which are used in the following way:

$$p_{1r}^{MSU} = min\left\{y_{1r}^{MSU}, y_{2r}^{MSU}\right\}, \tag{9}$$

$$p_{2r}^{MSU} = min\{p_{1r}^{MSU}, y_{3r}^{MSU}\},$$

$$\ldots$$

$$p_{(|I|-1)r}^{MSU} = min\left\{p_{(|I|-2)r}^{MSU}, y_{|I|r}^{MSU}\right\}.$$

In turn, each of these $(|I|-1)$ minimum operations are represented using six constraints in our optimization model.

To model each of the minimum operations (including two components), we also need one binary variable. We let binary help variable $s_{ir}^{MSU}, i \in \{1,\ldots,|I|-1\},\ r \in R$ be used in the i:th minimum operation in this sequence for subregion r.

The first of the minimum operations $p_{1r}^{MSU} = min\left\{y_{1r}^{MSU}, y_{2r}^{MSU}\right\}$, determining the minimum between $y_{1r}^{MSU}$and y_{2r}^{MSU} is modeled using the following (six) constraints.

$$y_{2r}^{MSU} - y_{1r}^{MSU} \leq Ms_{1r}^{MSU}, \tag{10}$$

$$y_{1r}^{MSU} - y_{2r}^{MSU} \leq M\left(1 - s_{1r}^{MSU}\right),$$

$$p_{1r}^{MSU} \leq y_{1r}^{MSU},$$

$$p_{1r}^{MSU} \leq y_{2r}^{MSU},$$

$$p_{1r}^{MSU} \geq y_{1r}^{MSU} - M\left(1 - s_{1r}^{MSU}\right),$$

$$p_{1r}^{MSU} \geq y_{2r}^{MSU} - Ms_{1r}^{MSU}.$$

The i:th $(2 \leq i \leq |I|-1)$ of the minimum operations $p_{ir}^{MSU} = min\left\{p_{(i-1)r}^{MSU}, y_{(i+1)r}^{MSU}\right\}$, determining the minimum between $p_{(i-1)r}^{MSU}$ and $y_{(i+1)r}^{MSU}$, is modeled using the following (six) constraints. Please note that there are in total $|I|-2$ such constraint sets for each subregion r.

$$y_{(i+1)r}^{MSU} - p_{(i-1)r}^{MSU} \leq Ms_{ir}^{MSU}, \tag{11}$$
$$p_{(i-1)r}^{MSU} - y_{(i+1)r}^{MSU} \leq M\left(1 - s_{ir}^{MSU}\right),$$
$$p_{ir}^{MSU} \leq p_{(i-1)r}^{MSU},$$
$$p_{ir}^{MSU} \leq y_{(i+1)r}^{MSU},$$
$$p_{ir}^{MSU} \geq p_{(i-1)r}^{MSU} - M(1 - s_{ir}^{MSU}),$$
$$p_{ir}^{MSU} \geq y_{(i+1)r}^{MSU} - Ms_{ir}^{MSU}.$$

Then, we use the constraint in Equation (12) (one for each $r \in R$) to acquire y_r^{MSU}. Please note that this constraint is needed in order to be consistent with the constraint set (11).

$$y_r^{MSU} = p_{(|I|-1)r}^{MSU}, \qquad r \in R. \tag{12}$$

The constraint $y_r = \min\{y_r^{MSU}, t_r^{RA}\}$ shown in Equation (5) captures the minimum value between y_r^{MSU} and t_r^{RA} for the patients located in subregion r. In the optimization model, this constraint is modeled using the following six constraints, where $v_r, r \in R$ is a binary help variable:

$$y_r^{MSU} - t_r^{RA} \leq Mv_r, \qquad r \in R \tag{13}$$
$$t_r^{RA} - y_r^{MSU} \leq M(1 - v_r),\ r \in R,$$
$$y_r \leq t_r^{RA},\ r \in R,$$
$$y_r \leq y_r^{MSU},\ r \in R,$$
$$y_r \geq t_r^{RA} - M(1 - v_r),\ r \in R,$$
$$y_r \geq y_r^{MSU} - Mv_r,\ r \in R.$$

The constraints $u^{max} \geq y_r$ and $u^{min} \leq y_r$ in Equations (6) and (7) capture the longest and shortest expected time to treatment for any subregion r and for any MSU in site i. The value of $u^{max} - u^{min}$ in the objective function, see Equation (2), refers to the range measure.

Finally, the constraint $\sum_{i \in I} x_i = N$ defined by Equation (8) specifies the number of MSUs to be placed in a region.

5. Scenario Study

In this section, we describe the application of our proposed optimization model to two counties in southern Sweden. We then describe the experimental results.

5.1 Scenario Description

To evaluate the efficacy of the presented optimization model, we apply it to the Blekinge and Kronoberg counties of Sweden, which are parts of Sweden's southern healthcare region (SHR). The SHR covers an area of 16,622 km^2 and encompasses four counties: Skåne, Blekinge, Halland, and Kronoberg. The SHR has 49 municipalities, and its population was 1,687,000 in 2018. In Sweden, over 21,000 stroke incidents occur annually, with 3,900 cases reported in SHR [17]. In SHR, there are 39 ambulance sites and 13 acute hospitals equipped with CT scanners. Using the standard solvers, for example, Gurobi, we realized that it would be difficult to solve the model for large problem instances, that is, the entire SHR. Therefore, we decided to test the model with two counties of SHR. Table 1 represents the demographic and geographic statistics for each county of SHR. Figure 1 shows an overview of SHR, where each green triangle (referred to by a specific circled number) and each purple circle corresponds to an ambulance site and an acute hospital, respectively. The borders of the Blekinge and Kronoberg counties are represented in red and blue, respectively. As shown in Figure 1, the ambulance sites in Blekinge are in Karlshamn (id: 15), Karlskrona (id: 16), Olofström (id: 26), Ronneby (id: 29), and Sölvesborg (id: 30), and ambulance sites in Kronoberg are in Älmhult (id: 1), Alvesta (id: 3), Lenhovda (id: 21), Lessebo (id: 22), Ljungby (id: 23), Markaryd (id: 25), Tingsryd (id: 34), and Växjö (id: 36).

Table 1. Demographic and geographic data of each county of SHR. NoM: number of municipalities; NoS: number of subregions; NoA: number of ambulance sites; and NoH: number of hospitals.

County	Population	NoM	NoS	NoA	NoH
Blekinge	134,188	5	1959	5	2
Halland	133,025	3	1603	4	1
Kronoberg	198,903	8	4233	8	2
Skåne	1,221,074	33	8827	22	8

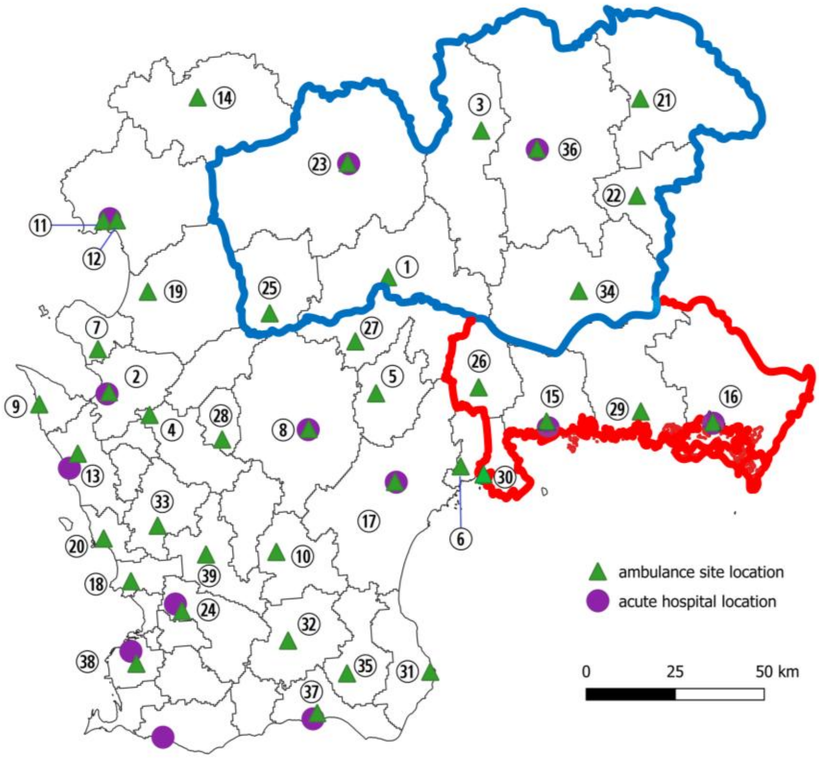


Figure 1. Overview of Sweden's southern healthcare (SHR). The purple circles and green triangles show the locations of acute hospitals and ambulance sites, respectively. The circled numbers indicate the corresponding ambulance site IDs. The borders of the Blekinge and Kronoberg counties are shown in red and blue, respectively.

We considered the same input data and assumptions as we did in our companion study [7]. In particular, we utilized the demographic data and stroke data for 2018 collected from Statistics Sweden [18] and Sweden's southern healthcare region committee [19], respectively. In our data, each county of SHR was divided into a set of non-overlapping subregions, each equaling to 1×1 km^2 and indicated by $r \in R$ so that the union of all subregions $\bigcup_{r \in R} r$ equals to the corresponding county of SHR. The demographic data included the number of inhabitants for each subregion $r \in R$ and each of the 21 assumed age groups, that is $\{[0,4), [4,8), \dots, [95,99), [100, \infty)\}$. In addition, the stroke data included the number of stroke cases for each age group in each county of SHR. Using the provided data, we calculated Q_r, indicated in Section 4, for each subregion $r \in R$, obtained by dividing the expected number of stroke cases in subregion r by the total expected number of stroke cases in the SHR.

In the scenario study, we aimed to identify the optimal locations of different numbers of MSUs in either Blekinge or Kronoberg using the proposed optimization model. In the experiments, we took into consideration that every ambulance site within the region could potentially serve as a location for placing an MSU. In addition, for all experiments, we measured the results using only the expected time to treatment. We also compared the experimental results of placing MSU(s) with the experimental results of the baseline, representing the current situation in the SHR, where there are only regular ambulances across all 39 ambulance sites in the SHR.

We solved the problem in Gurobi[1] 10.0.0, which uses the barrier and simplex algorithms to solve continuous relaxations of mixed-integer models and continuous models. In all experiments, we solved the described problem using the barrier and simplex algorithms. All of the code was written in Jupyter Notebook using Python on a computer with 32-gigabyte memory (RAM) and a Core(TM) i7-8650U CPU 1.90 gigahertz Intel(R) processor.

5.2 Experimental Results

As mentioned above, we applied our model to different parts of the SHR, that is, Blekinge and Kronoberg counties. To demonstrate the functionality of our optimization model and to explore how large problem instances can be solved using this approach, we initially tried to apply it to a smaller county of SHR, that is, Blekinge, which is a smaller region and which has a lower number of ambulance sites compared to the entire SHR. We, then, applied the model to a broader region, that is, Kroboberg county, with a higher number of ambulance sites. The reason that we did not represent the application of the model to the complete SHR, which is a large area, is that it was challenging and time-consuming to optimally solve our proposed model for such a large region with the corresponding large amount of input data.

The experimental results for the Blekinge county are presented in Table 2. We considered two situations regarding the number of ambulance sites that are available for placing MSUs in Blekinge: Situation 1) the number of available ambulance sites corresponds to the number of ambulance sites in Blekinge;

[1] Available: https://www.gurobi.com

Table 2. Experimental results for the Blekinge county. NoAAS: number of available ambulance sites for placing MSUs; NoM: number of MSUs to place in the county; Alg.: algorithm used to solve the problem; MSU IDs: found optimal MSU site IDs, denoted by numbers within the square brackets; Ex. Time: Execution time (in seconds); Tr.: objective function corresponding to tradeoff value (in hour); Ra.: range (in hours); ATT: average time to treatment (in hour); WATT: weighted average time to treatment (in hour); ES: exhaustive search; S: simplex; B: barrier, and S&B: simplex and barrier.

NoAAS	NoM	Alg.	MSU IDs	Ex. Time	Tr.	Ra.	WATT	ATT
Baseline	-	-	-	-	1.39	1.44	1.34	1.61
5 & 11	1	ES	[29]	-	1.09	1.16	1.01	1.11
5	1	S&B	[29]	22	1.09	1.16	1.01	1.11
7 (11)	1	S (B)	[29]	42 (45)	1.09	1.16	1.01	1.11
5 & 11	2	ES	[15,16]	-	0.87	0.89	0.84	0.97
5	2	S&B	[15,16]	572	0.87	0.89	0.84	0.97
7 (11)	2	S (B)	[15,16]	978 (1008)	0.87	0.89	0.84	0.97
5	3	ES	[15,16,29]	-	0.81	0.83	0.79	0.93
5	3	S&B	[15,16,29]	466	0.81	0.83	0.79	0.93
7 (11)	3	S (B)	[15,16,34]	1243 (1127)	0.82	0.79	0.84	0.95
11	3	ES	[15,16,34]	-	0.82	0.79	0.84	0.95
5	4	ES	[15,16,26,29]	-	0.78	0.81	0.75	0.89
5	4	S&B	[15,16,26,29]	24	0.78	0.81	0.75	0.89
7 (11)	4	S (B)	[15,16,26,34]	1003 (995)	0.81	0.81	0.80	0.91
11	4	ES	[15,16,26,34]	-	0.81	0.81	0.80	0.91

Situation 2) the number of available ambulance sites corresponds to the number of ambulance sites in Blekinge + all ambulance sites located in the neighborhood of Blekinge. As can be seen the Figure 1, there are six ambulance sites close to Blekinge, where the two nearest ambulance sites are Bromölla (id: 6) and Tingsryd (id: 34).

Since there are 5 ambulance sites in the Blekinge county, in the experiments, we solved the problem for placing 1 to 4 MSUs. According to Table 2, by adding the number of MSUs, the tradeoff value decreases in comparison with the baseline (where there is no MSU in Blekinge). The results also demonstrate that by using MSU(s) in Blekinge, the values of the tradeoff, WATT, range, and average time to treatment are expected to decrease compared to the baseline. In particular, by placing two MSUs in Blekinge, it is possible to make the treatment available within an hour for all inhabitants living in Blekinge.

When we solved the problem considering Situation 1 (only ambulance sites in Blekinge), the Simplex and Barrier algorithms produced the same results for each

MSU placement. For Situation 1, we pointed out the minimum execution time between Simplex and Barrier in Table 2.

For Situation 2, where we considered 11 ambulance sites (5 ambulance sites in Blekinge and 6 neighborhood ambulance sites), the Gurobi solver using the Simplex algorithm had difficulty in solving the problem for placing of 1 to 4 MSUs. We instead decided to only consider the 2 nearest ambulance sites to the existing ambulance sites of Blekinge and perform the experiments with 7 ambulance sites, shown in the third row of each MSU placement in Table 2. However, using the Barrier algorithm, Gurobi could solve the problem considering 11 ambulance sites in a feasible amount of time. According to the Gurobi documentation[2], the reason is probably that the Barrier algorithm is more efficient for complex models with large size. The results of the Barrier algorithm for Situation 2 are presented in parentheses in the third row for each MSU placement.

In Table 2, the comparison of the results of Situation 1 and Situation 2 shows that the identified MSU locations are equal when placing 1 and 2 MSUs. However, the identified MSU locations are different when placing 3 and 4 MSUs, where the corresponding tradeoff values of Situation 1 are smaller than Situation 2.

According to Table 2, in Situation 1 and Situation 2, the highest execution times are recorded when placing 2 MSUs (572 seconds) and 3 MSUs (1243 seconds for Simplex and 1127 seconds for Barrier), respectively.

In order to verify the optimal solutions and optimal objective function values obtained using our optimization model, we compared the output of our model with the exhaustive search, proposed in our companion paper [7], for placing 1, 2, 3, and 4 MSUs in Blekinge, presented in Table 2. In all MSU placements and Situations, the identified solutions and objective function values are the same both for our proposed model and for the exhaustive search.

In Table 3, we present the results of applying our model to the Kronberg county. In the experiments, we assumed that only ambulance sites in Kronberg can be used for placing MSUs. Considering the complexity of solving the model, we, further, assumed that it is relevant to solve the problem of placing 1 to 5 MSUs in Kronoberg.

According to Table 3, by adding the number of MSUs, the tradeoff value decreases in comparison with the baseline (where there is no MSU in Kronoberg). In Table 3, the results also show that by placing MSU(s) in Kronoberg, the values

[2] Available: https://www.gurobi.com/documentation/

Table 3. Experimental results for the Kronoberg county. The abbreviations are the same as in Table 2.

NoAAS	NoM	Alg.	MSU IDs	Ex. Time	Tr.	Ra.	WATT	ATT
Baseline	-	-	-	-	1.65	1.84	1.45	1.78
8	1	ES	[3]	-	1.32	1.53	1.11	1.31
8	1	S (B)	[3]	2195 (133)	1.32	1.53	1.11	1.31
8	2	ES	[23,36]	-	1.09	1.24	0.94	1.13
8	2	S	[23,36]	2667	1.09	1.24	0.94	1.13
8	3	ES	[21,23,36]	-	1.01	1.10	0.91	1.10
8	3	S	[21,23,36]	3481	1.01	1.10	0.91	1.10
8	4	ES	[21,23,34,36]	-	0.99	1.10	0.87	1.02
8	4	S	[21,23,34,36]	4099	0.99	1.10	0.87	1.02
8	5	ES	[21,23,25,34,36]	-	0.97	1.10	0.83	0.98
8	5	S	[21,23,25,34,36]	2340	0.97	1.10	0.83	0.98

of the tradeoff, WATT, range, and average time to treatment are expected to reduce compared to the baseline. Especially, placing 5 MSUs in Kronoberg would potentially provide treatment within an hour for all inhabitants living there.

It can be observed in Table 3 that when we solve the problem of placing one MSU in Kronoberg, the Simplex and Barrier algorithms produce the same results. However, when placing more than one MSU in Kronoberg, the Barrier algorithm had difficulties in finding feasible solutions for placing 2 to 5 MSUs. Alternatively, using the Simplex algorithm, Gurobi could solve the problem for different numbers of MSUs within a feasible amount of time. As mentioned above, the reason appears to be that the Barrier algorithm tends to be quicker when handling large complex models, but it exhibits greater numerical sensitivity. On the other hand, the simplex algorithm is generally less affected by numerical issues. According to Table 3, the highest execution time (4099 seconds) is recorded when placing 4 MSUs in Kronoberg.

Similar to Table 2, we compared the output of the presented optimization model with the exhaustive search for placing different numbers of MSUs in Kronoberg, presented in Table 3. As can be seen for all MSU placements, the identified solutions and objective function values are the same both for our proposed model and for the exhaustive search.

From the conducted experiments, we could explore to what extent large problem instances can be solved using our optimization model, and in that way, we could learn about the limits of using the Gurobi solver to solve the described problem.

6. Conclusions

We have presented a MILP model for the optimal placement of MSUs in a geographic region. The objective function of our optimization model is a tradeoff function proposed in our prior study [7], used to tradeoff the equity and efficiency perspectives for the MSU placement problem while aiming to provide shorter time to treatment and equal service for residents living in a region. To evaluate our optimization model, we conducted a scenario study to place MSUs in the Blekinge and Kronoberg counties of Sweden. Applying the model to smaller counties provided us the opportunity to assess the model's functionality and performance on a more manageable scale before scaling it up to larger problem instances. In the presented model, the time needed to identify an optimal solution for the given problem instances indicated the complexity of the MSU placement problem. The experimental results, supported by the results of the exhaustive search approach presented in previous research [7], indicated that the proposed optimization model is able to find the optimal MSU locations concerning the defined objective function and constraints. The results of the experiments also showed that using our proposed optimization model for the MSU placement problem enabled to cut down the expected time to treatment for most residents compared to the baseline. From the experimental results, we concluded that by placing 2 and 5 MSUs in Blekinge and Kronoberg, respectively, it is likely to achieve access to treatment within an hour for all inhabitants living there, which is often considered an important goal.

The focus of the current paper was on validating the correctness of the proposed optimization model and illustrating the possible idea of placing MSU(s) in a region using the tradeoff perspective. As mentioned above, solving large problem instances, for example, the SHR, is computationally expensive, in particular for standard optimization solvers. For future work, we plan to investigate the use of heuristics to solve large problem instances within a reasonable time frame.

References

[1] World Stroke Organization, “Facts and figures about stroke,” [Online]. Available: https://www.world-stroke.org/world-stroke-day-campaign/whystroke-matters/learn-about-stroke. [Accessed: Dec. 20, 2019].

[2] M. Ebinger, B. Winter, M. Wendt, J. E. Weber, C. Waldschmidt, M. Rozanski, A. Kunz, P. Koch, P. A. Kellner, and D. Gierhake, “Effect of the use of ambulance-based thrombolysis on time to thrombolysis in acute ischemic stroke: a randomized clinical trial,” *Jama*, vol. 311, no. 16, pp. 1622-1631, 2014.

[3] T. G. Phan, R. Beare, V. Srikanth, and H. Ma, "Googling location for Mobile Stroke Unit hub in metropolitan Sydney," *Frontiers in neurology*, vol. 10, p. 810, 2019.

[4] J. P. Rhudy Jr, A. W. Alexandrov, J. Rike, T. Bryndziar, A. H. Z. Maleki, V. Swatzell, W. Dusenbury, E. J. Metter, and A. V. Alexandrov, "Geospatial visualization of mobile stroke unit dispatches: a method to optimize service performance," *Interventional neurology*, vol. 7, no. 6, pp. 464-470, 2018.

[5] S. Mathur, S. Walter, I. Q. Grunwald, S. A. Helwig, M. Lesmeister, and K. Fassbender, "Improving prehospital stroke services in rural and underserved settings with mobile stroke units," *Frontiers in Neurology*, vol. 10, 2019.

[6] O. Dahllöf, F. Hofwimmer, J. Holmgren, and J. Petersson, "Optimal placement of Mobile Stroke Units considering the perspectives of equality and efficiency," in *Proceedings of the 8th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH 2018)*, vol. 141, pp. 311-318, 2018.

[7] S. Amouzad Mahdiraji, J. Holmgren, R. C. Mihailescu, and J. Petersson, "An optimization model for the tradeoff between efficiency and equity for mobile stroke unit placement," In *Innovation in Medicine and Healthcare: Proceedings of 9th KES International Conference on Innovation in Medicine and Healthcare (KES-InMed-21)*, pp. 183-193, 2021.

[8] V. Bélanger, A. Ruiz, and P. Soriano, "Recent optimization models and trends in location, relocation, and dispatching of emergency medical vehicles," *European Journal of Operational Research*, vol. 272, no. 1, pp. 1-23, 2019.

[9] Tavakoli and C. Lightner, "Implementing a mathematical model for locating EMS vehicles in Fayetteville, NC.," *Computers & Operations Research*, vol. 31, no. 9, pp. 1549-1563, 2004.

[10] J. Røislien, P. L. van den Berg, T. Lindner, E. Zakariassen, K. Aardal, and J. T. van Essen, "Exploring optimal air ambulance base locations in Norway using advanced mathematical modelling," *Injury prevention*, vol. 23, no. 1, pp.10-15, 2017.

[11] H. Leknes, E. S. Aartun, H. Andersson, M. Christiansen, and T. A. Granberg, "Strategic ambulance location for heterogeneous regions," *European Journal of Operational Research*, vol. 260. No. 1, pp. 122-133, 2017.

[12] M. A. Akdoğan, Z. P. Bayındır, C. Iyigun, "Locating emergency vehicles with an approximate queuing model and a meta-heuristic solution approach," *Transportation Research Part C: Emerging Technologies*, vol. 90, pp. 134-155, 2018.

[13] T. Tlili, S. Abidi, and S. Krichen, "A mathematical model for efficient emergency transportation in a disaster situation," *The American Journal of Emergency Medicine*, vol. 36, no. 9, pp. 1585-1590, 2018.

[14] J. A. Acuna, J. L. Zayas-Castro, and H. Charkhgard, "Ambulance allocation optimization model for the overcrowding problem in US emergency departments: A case study in Florida," *Socio-Economic Planning Sciences*, vol. 71, p. 100747, 2020.

[15] S.-p. Wan, Z.-h. Chen, and J.-y. Dong, "Bi-objective trapezoidal fuzzy mixed integer linear program-based distribution center location decision for large-scale emergencies," *Applied Soft Computing*, vol. 110, p. 107757, 2021.

[16] S. Amouzad Mahdiraji, O. Dahllöf, F. Hofwimmer, J. Holmgren, R.-C. Mihailescu, and J. Petersson, "Mobile stroke units for acute stroke care in the south of Sweden," *Cogent Engineering*, vol. 8, no. 1, 2021, doi: https://doi.org/10.1080/23311916.2021.1874084.

[17] The Swedish Stroke Register, "Stroke registrations," [Online]. Available: https://www.riksstroke.org/sve/forskning-statistikoch-verksamhetsutveckling/statistik/registreringar. [Accessed: Dec. 20, 2019].

[18] Statistics Sweden, “demographic data 2018,” [Online]. Available: https://www.scb.se. [Accessed: Jul. 10, 2018].

[19] Sweden’s Southern Regional Health Care Committee, “stroke data 2018,” [Online]. Available: https://sodrasjukvardsregionen.se. [Accessed: Jul. 10, 2018].

PAPER VI - SIMULATION-BASED ANALYSIS OF CO-DISPATCHING IN PREHOSPITAL STROKE CARE

Saeid Amouzad Mahdiraji, Johan Holmgren, Radu-Casian Mihailescu, and Jesper Petersson

ABSTRACT

A mobile stroke unit (MSU) is a specialized ambulance, enabling to shorten the time to diagnosis and treatment for stroke patients. In the current paper, we present a simulation-based approach to study the potential impacts of collaborative use of regular ambulances and MSUs in prehospital transportation for stroke patients, denoted as co-dispatching. We integrated a co-dispatch policy in an existing modeling framework for constructing emergency medical services simulation models. In a case study, we applied the extended framework to southern Sweden to evaluate the effectiveness of using the co-dispatch policy for different types of stroke. The results indicate reduced time to diagnosis and treatment for stroke patients when using the co-dispatch policy compared to the situation where either a regular ambulance or an MSU is assigned for a stroke incident.

PUBLISHED IN

Proceedings of the 15th International Conference on Ambient Systems, Networks and Technologies (ANT 2024), pp. 412-419, 2024.

1. Introduction

Stroke is a global leading cause of disability and mortality, resulting from interrupted or reduced blood supply to the brain [1]. There are three common types of stroke: ischemic, hemorrhagic, and transient ischemic attack (TIA), where each type requires specific treatment. Accurate diagnosis of a stroke requires a computed tomography (CT) scan on the patient's brain, and this technique is typically only accessible at acute hospitals. An ischemic stroke, the most common stroke type, occurs when one or more blood clots reduce the blood flow to the brain, and the patient can be treated through thrombolysis and, in certain cases, through thrombectomy, which is only available in special clinics (SCs). Hemorrhagic strokes happen when a blood vessel in the brain ruptures, requiring an early start of blood pressure-lowering therapy. TIA is a temporary disruption of blood flow to the brain, allowing for complete recovery of brain function. Timely intervention is crucial for the successful recovery of stroke patients, emphasizing the need for rapid and effective prehospital transportation and care [2].

In emergency medical services (EMS), a single dispatch policy refers to assigning only one type of emergency vehicle (EV) for an incident. Figure 1. A shows a single dispatch policy using a regular ambulance (RA), where the patient can be transported to the closest hospital or an SC. Additionally, the patient sometimes needs to be transported to the secondary hospital (that is, the SC), as shown in Figure 1. B. The limitations of using RA-dispatch in prehospital stroke care have led to the exploration of alternative policies, such as deploying mobile stroke units (MSUs). MSUs are specialized ambulances equipped with a CT scanner, enabling ambulance personnel to perform diagnosis and intravenous stroke treatment (that is, thrombolysis) inside the MSU. Therefore, the use of MSUs can potentially reduce the time to treatment by eliminating the time required for transporting the patient to the hospital.

Several studies aim to improve prehospital stroke care through simulation and optimization modeling. In particular, some studies use simulation to assess stroke transportation policies, including MSUs [3-5]. For example, Amouzad Mahdiraji et al. [3] present a discrete event simulation (DES) model to assess stroke transportation policies, particularly those involving MSUs. The authors further build a modeling framework applicable to various EMS simulation models for different medical conditions, including stroke [4].

In EMS, co-dispatching or a rendezvous approach refers to a situation where two or more EVs are simultaneously dispatched to an emergency incident and collaborate until the patient receives appropriate treatment. For example, involving both an RA and an MSU in the transportation of a suspected stroke patient can provide more timely and effective care [6]. In co-dispatching, upon receiving a stroke-related call, one EV, such as an RA, is dispatched to the patient's location, while the other EV, such as an MSU, is sent directly to a predetermined meeting point. After picking up the stroke patient, the RA travels towards the meeting point, where the two EVs come together to transfer the patient. Then, the MSU transports the patient to an appropriate hospital based on the diagnosis made inside the MSU. Alternatively, the MSU may be dispatched to the patient's location, and the RA to the meeting point, depending on operational considerations. Figure 1. D illustrates co-dispatching, where both an RA and an MSU are dispatched simultaneously to the patient's location (that is, the meeting point that is assumed in this paper). Then, depending on the diagnosis outcome, the patient is transported to either an acute hospital or an SC using one of the involved EVs (that is, either the RA or the MSU).

Previous research has delved into various aspects of prehospital stroke patient transportation; however, to the best of our knowledge, a simulation-based evaluation of co-dispatching remains unexplored. In the current paper, we present a simulation-based analysis of co-dispatching to investigate the potential advantages associated with the collaborative use of RAs and MSUs in prehospital stroke care for different types of stroke. The contribution of this paper is two-fold: i) an extension of the modeling framework introduced in our companion study [4], by incorporating the co-dispatch policy, and ii) the application of the framework to southern Sweden to study the benefits of using the co-dispatch policy for stroke patients with different types of stroke.

The remainder of the paper is structured as follows. In Section 2, we present how the co-dispatch policy is integrated into the modeling framework proposed in our previous study [4]. Section 3 describes the scenario study, followed by the results and a discussion. Finally, Section 4 concludes the paper.

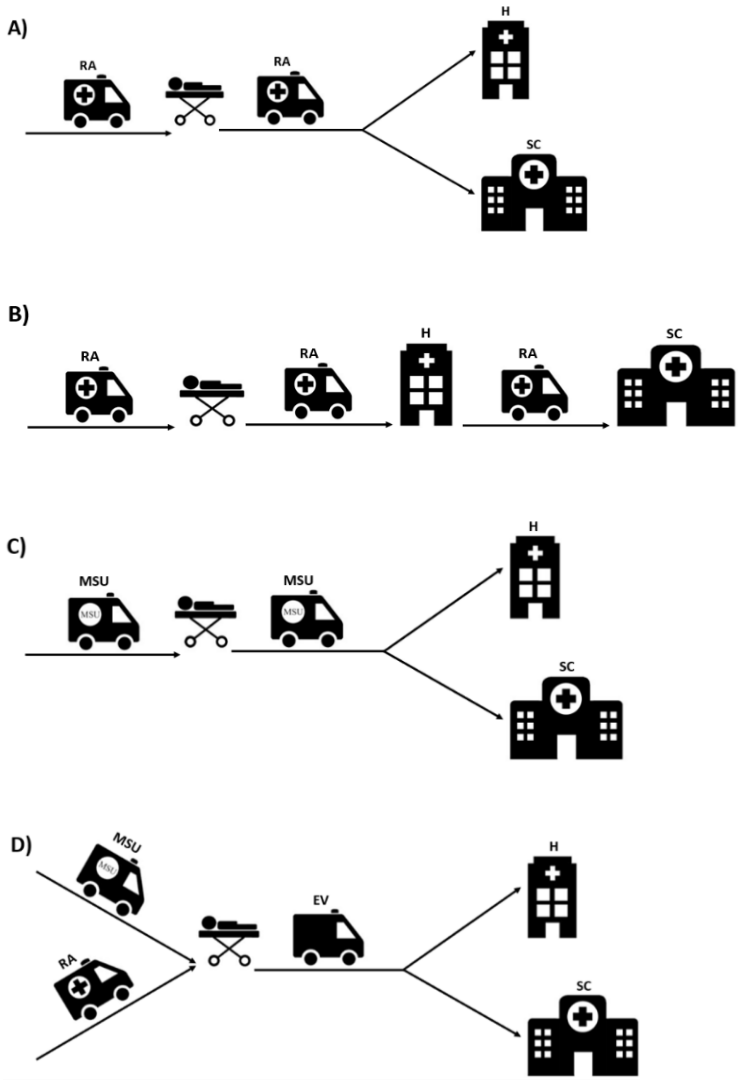


Figure 1. Different policies made in prehospital stroke care for an acute stroke patient. A) Single dispatch policy using an RA, where the patient is transported to the closest medical center, that is, either the closest hospital or a special clinic (SC); B) Single dispatch policy using an RA, where the patient is transported to the closest hospital and then further transported to the SC; C) Single dispatch policy using an MSU, where the patient is transported to either the closest hospital or an SC considering the decision made after patient diagnosis inside the MSU or distance to the SC; D) Co-dispatching, where both an RA and an MSU are dispatched simultaneously to the patient's location, and then based on the outcome of patient diagnosis inside the MSU, the patient will be transported to the either closest hospital or an SC using an EV (that is, either an RA or an MSU). Graphs in this figure are inspired by Mathur et al. [6].

2. A Co-dispatch Policy in the Modeling Framework

In a prior study [4], we proposed a modeling framework to simplify the construction of EMS simulation models. The framework includes model construction and execution phases. During the construction phase, the framework uses various input, including geographical data, patient data, ambulance data, hospital data, and a care chain specification, to build a DES model for a specific EMS scenario. In the execution phase, the created model is executed using the input data, and the given care chain specification is applied to each simulated patient. The framework utilizes building blocks to construct care chains of EMS activities as flowcharts, where each block represents a generic activity, either a decision-making activity or a pure task activity (see Figures 2-4). The original version of the framework supports only a single dispatch policy, assigning either an RA or an MSU for a stroke incident. In the current paper, we extend the framework to include the co-dispatch policy.

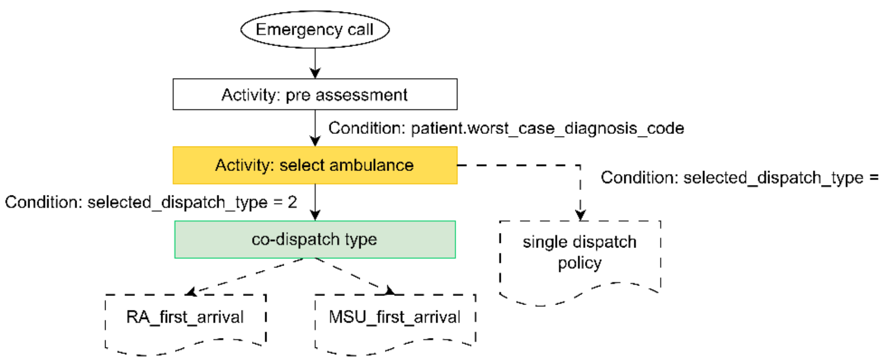


Figure 2. Overview of the proposed EMS care chain activities, including co-dispatch policy, from emergency call until the co-dispatch type activity. The proposed chain of EMS activities for the single dispatch policy is presented in our prior study [4]. The continuation of the two branches of the co-dispatch type activity, RA_first_arrival and MSU_first_arrival, are illustrated in Figure 3 and Figure 4, respectively.

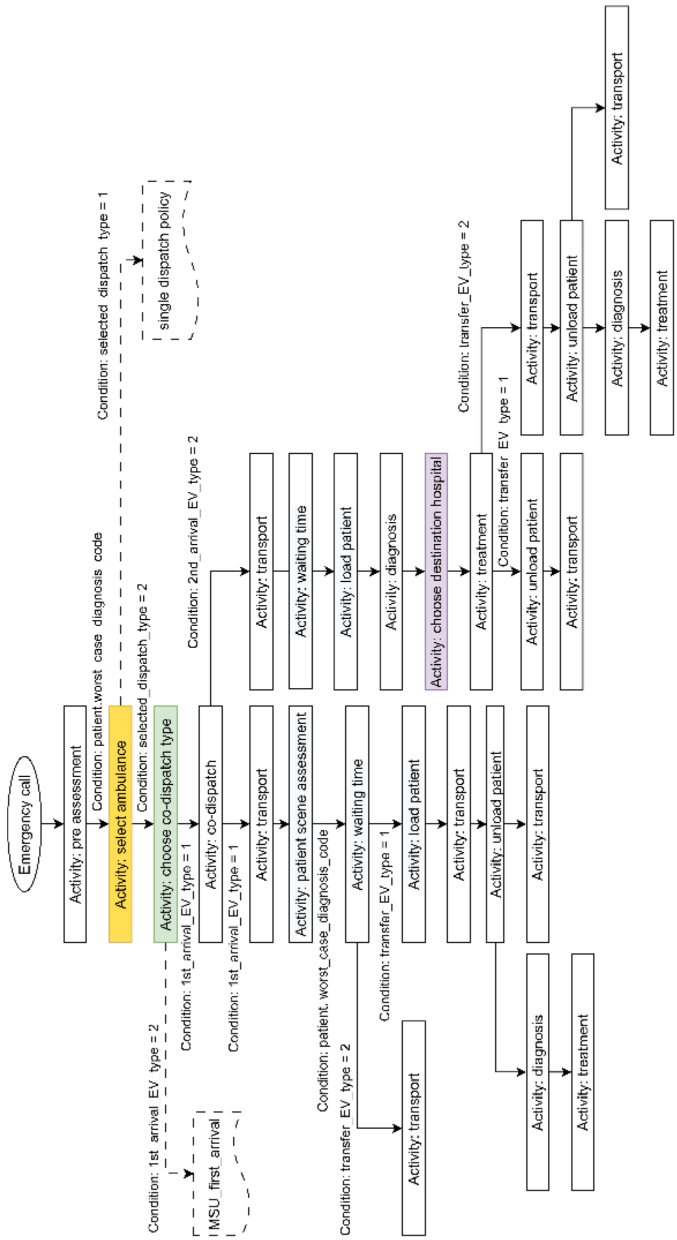


Figure 3. Continuation of Figure 2; proposed care chain activities of the co-dispatch policy for a suspected stroke patient, where an RA arrives earlier than an MSU at the patient's location. The colored boxes demonstrate the decision-making activities.

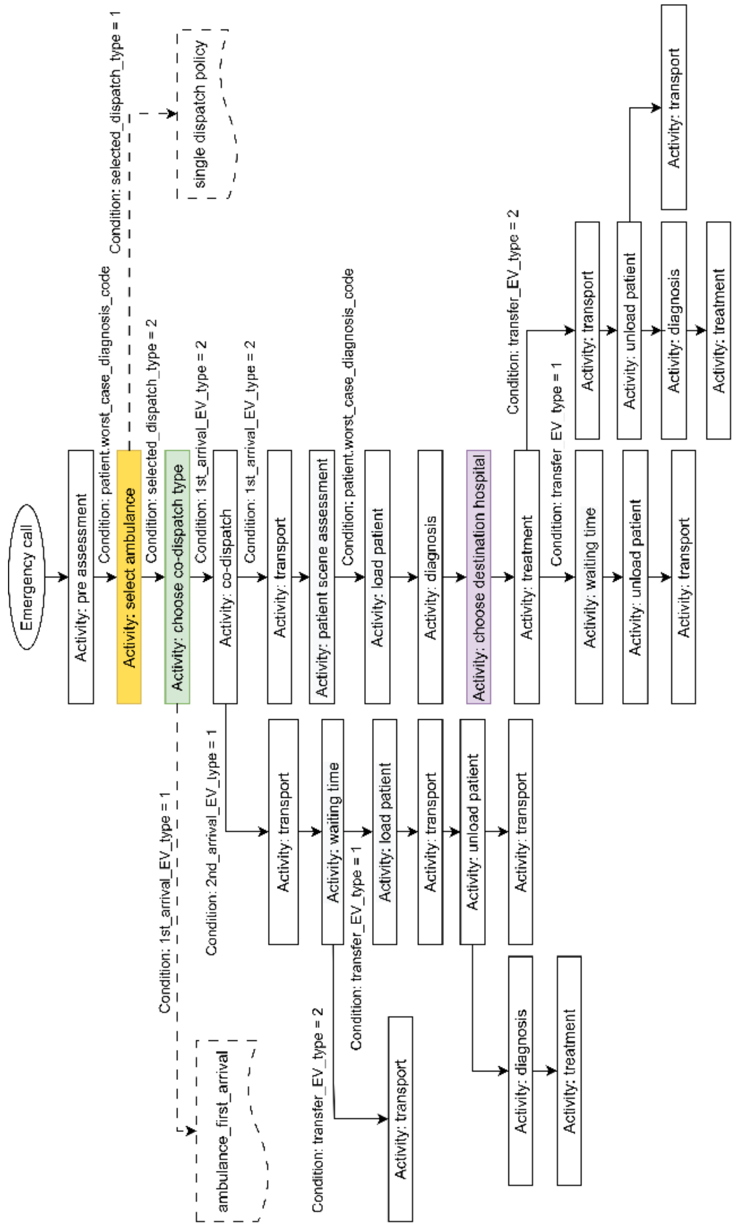


Figure 4. Continuation of Figure 2; proposed care chain activities of the co-dispatch policy for a suspected stroke patient, where an MSU arrives earlier than an RA at the patient's location. The colored boxes demonstrate the decision-making activities.

In Figures 2-4, we illustrate the chain of EMS activities for a suspected stroke patient, starting from when the emergency center receives a call concerning an incident until the patient receives treatment. As shown in Figure 2, following the execution of the *select ambulance* activity, the care chain advances to either the single dispatch or the co-dispatch policy. In the co-dispatching branch, the *co-dispatch type* activity is split into two branches: *RA_first_arrival* (Figure 3) and *MSU_first_arrival* (Figure 4), implying the situations where the RA arrives before the MSU at the patient's location and the MSU arrives before the RA at the patient's location, respectively. The activities for the co-dispatch policy are split into two distinct branches as the sequence of activities are different depending on whether an RA or an MSU arrives first at the patient's location. A detailed presentation of the EMS care chain activities for the single dispatch policy is provided in our prior study [4].

The key decisions modeled in the framework, and which are included in the care chain example in Figures 2-4 are choosing 1) the type of EV to be dispatched towards the patient, 2) the dispatch type between single dispatch and co-dispatch, 3) the destination hospital. To reflect these decisions, we included a range of policies for the selection of ambulances, dispatch types, and hospitals, described in the following.

The ambulance selection policies aim to choose one or two EVs and their corresponding types by minimizing either the expected time to 1) the patient, 2) the diagnosis, or 3) the initiation of treatment. The user decides which of these policies should be chosen for the ambulance selection process. The values of the time to diagnosis and time to treatment are estimated based on the identified patient's symptoms over the phone during the preassessment activity.

Dispatch policies have a tied link with ambulance selection policies, where we decide whether a single dispatch policy (dispatching either an RA or an MSU) or a co-dispatch policy should be chosen for a suspected stroke incident. The dispatch policy is designed to choose the dispatch type (either single dispatch or co-dispatch) that minimizes the expected time according to the chosen ambulance selection policy (that is, either *time to patient*, *time to diagnosis*, or *time to treatment*). The supported policies are:

- **Single dispatch**: Either an RA or an MSU is assigned for a stroke incident (see Figure 1. A to Figure 1. C).
- **Co-dispatch**: Two EVs collaborate in prehospital stroke care and transportation (see Figure 1. D). In this policy, when a stroke incident occurs and both an RA and an MSU are dispatched to the meeting point, that is, the patient's location, two situations may occur:

- **Ambulance_first_arrival**: (see Figure 3) If the RA arrives earlier than the MSU at the patient's location, the RA staff performs patient scene assessment (and waits) until the MSU arrives. When the MSU arrives at the patient's location, the patient is loaded into the MSU, and the MSU staff performs diagnosis and treatment. After the diagnosis procedure and choosing a destination hospital, the patient is transported either by the MSU or the RA to the destination hospital. When the transferring EV starts to drive to the chosen hospital, the other EV is dismissed and returns to its station. Once the transferring EV has unloaded the patient at the hospital, it returns to its station.
- **MSU_first_arrival**: (see Figure 4) If the MSU arrives earlier than the RA to the patient's location, the MSU staff performs patient scene assessment, loads the patient to the MSU, performs diagnosis and treatment, and chooses a destination hospital. Upon the arrival of the RA at the patient's location, the patient is transported either by the MSU or the RA to the destination hospital. Simultaneously, the other EV is released and heads back to its station. Once the transferring EV has unloaded the patient at the hospital, it returns to its station.

The use of co-dispatching, including an RA and an MSU, in prehospital stroke care offers several advantages for stroke patients. For example, in the ambulance_first_arrival situation, once the RA arrives at the patient's location, its staff can perform the patient scene assessment, enabling the patient to be diagnosed inside the MSU upon arrival. Additionally, in the MSU_first_arrival situation, if the RA transports the patient to the chosen hospital, the MSU becomes available immediately for subsequent emergency operations.

In the co-dispatch policy, we define two distinct policies, referred to as transferring to hospital policies, to decide which of the assigned EVs in co-dispatching should transport a stroke patient to the hospital. The choice of the EV for the patient transfer is based on the user preference between the following two transferring to hospital policies:

- **Faster vehicle**: The EV that can transfer the patient fastest to the destination hospital is chosen.
- **Two vehicles collaboration**: Refers to the collaboration of both assigned EVs, meaning that if one EV arrives earlier at the patient's location, the other EV will be assigned to transfer the patient to the chosen hospital.

Hospital selection policies determine to which hospital the patient should be transported: 1) directly to the closest hospital, 2) directly to the SC, or 3) to either the closest hospital or the SC depending on the patient scene assessment (and diagnosis) and how far the SC is from the patient's location, phrased as fusion. As an illustration of the fusion policy, following the diagnosis of the patient inside the MSU, if specific treatment is deemed necessary, the patient is transported to the SC; otherwise, the patient is transported to the nearest hospital. Using RA-dispatch, if the patient requires specific treatment, yet the SC is far away, the patient is first transported to the closest hospital; then, if it is needed, the patient will be transported to the SC using the same RA (see Figure 1. B).

By conducting a series of simulation runs, it is possible to compare different decision policies to observe from which policy the patients are expected to benefit most. It is worth noting that in the MSU-dispatch or co-dispatch policy, the framework settings allow us to exclude diagnosis and treatment inside the MSU depending on the patient's symptoms, diagnosis inside the MSU, and preferred treatment.

In Figures 2-4, the decision to select dispatch policy and select EV(s) and their type are made in the *select ambulance* activity, shown in orange. In the co-dispatch policy, the type of EV that should transfer the patient to the hospital (transferring to hospital policy) is determined in the *choose co-dispatch type* activity, shown in green. Moreover, the decision for choosing the destination hospital is made in the *choose destination hospital* activity, colored by purple, using the predefined hospital selection policy.

3. Scenario Study

We evaluated our approach using Sweden's Southern Healthcare Region (SHR), which is depicted in Figure 5, where the ambulance sites and acute hospitals are shown by green triangles and purple circles, respectively, and each ambulance site has its unique ID. Currently, no MSU operates in the SHR. Lund SUS is the only SC for thrombectomy treatment, located near ambulance site ID: 24. According to the Swedish Stroke Register (Riksstroke) [7], Sweden recorded over 20,000 annual stroke incidents in 2022, with 3,900 cases reported in the SHR. Of the registered stroke cases in Sweden, 87% were diagnosed with ischemic stroke, and 13% with hemorrhagic stroke. Among the ischemic stroke cases, 81% did not receive treatment, while the remaining 19% underwent recanalization therapy, including thrombolysis alone (11%), thrombolysis combined with thrombectomy (3%), and thrombectomy alone (5%).

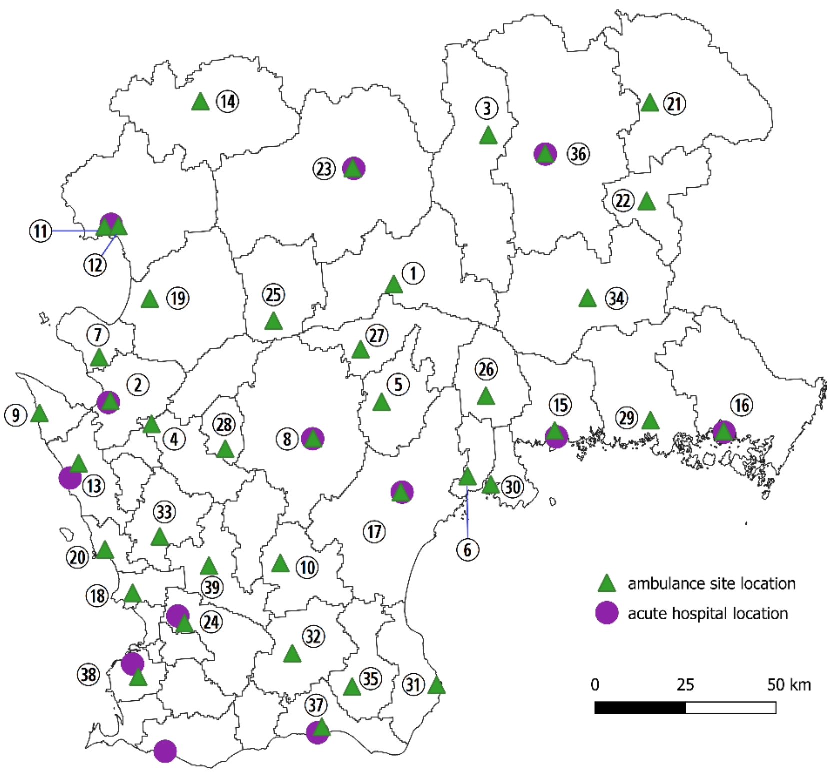


Figure 5. Overview of Sweden's southern healthcare Sweden region (SHR), reproduced from [8]. The purple circles and green triangles show the locations of acute hospitals and ambulance sites, respectively. The circled numbers indicate the corresponding ambulance site IDs.

We built a model for the EMS care chain of stroke patients in the SHR and carried out a series of simulation runs. In the simulation runs, we considered all decision policies described in Section 2. We used the same input data, MSU locations, and assumptions as we did in our companion study [4]. The geographic region, that is, the SHR, is divided into disjoint 1×1 km^2 subregions, which are used to locate patients, hospitals, and ambulance sites. The stroke patient data, created using a Poisson distribution, is a synthetic population distributed over SHR (details in a previous study [3]). Each patient in the population has specific attributes such as incident time, location, age, symptoms, diagnoses, and preferred treatment. We assumed that all generated patients were acute stroke patients who require diagnosis, and potentially, treatment either inside the MSU or at a hospital. The TIA cases were not considered. We also assumed ischemic stroke cases can receive both diagnosis and treatment (thrombolysis) at the patient's location. However, the patient will be transported to the hospital for (further) diagnosis and treatment (see Figure 1. C). In all situations, the time to diagnosis/treatment is the expected time from when a stroke occurs until the patient receives diagnosis/treatment either inside the MSU or at the hospital.

We considered three scenarios aligned with the defined dispatch policies: 1) single dispatch using RA only (current situation in the SHR), 2) single dispatch using MSU only, and 3) co-dispatch. We evaluated these scenarios concerning the other mentioned decision policies, namely ambulance selection, hospital selection, and transferring to hospital policies. In the experiments, the MSU locations in the SHR correspond to the findings from our previous paper [8], where we presented an optimization model for the MSU placement problem to make a balance between maximizing population coverage and ensuring equitable service. Thus, we conducted experiments by placing one MSU in Alvesta (ID: 3) or three MSUs in Ängelholm (ID: 2), Alvesta (ID: 3), and Malmö (ID: 38).

In the co-dispatch policy, the *waiting time* activity determines how long each EV waits at the patient's location considering each of the mentioned co-dispatch types, which corresponds to one of the following situations:

- For the ambulance_first_arrival situation (RA arrives earlier, MSU arrives later):
    - When the RA arrives earlier at the patient's location and its paramedics have done the patient scene assessment, the RA should either wait until the MSU arrives to load the patient to the MSU or wait until the end of diagnosis and treatment inside the MSU to transfer the patient to the chosen hospital.
    - When the MSU arrives at the patient's location, it should wait until the RA's paramedics perform the patient scene assessment. Then, the patient is loaded into the MSU for diagnosis and treatment.
- For the MSU_first_arrival situation (MSU arrives earlier, RA arrives later):
    - When the MSU arrives at the patient's location, after performing assessment, diagnosis, and treatment procedure, it should wait until the RA arrives to transport the patient to the chosen hospital.

- When the RA arrives at the patient's location, it should wait until the patient receives diagnosis and treatment inside the MSU before transporting the patient to the chosen hospital.

3.1 Experimental Results

In the experiments, we employed various policies including time to diagnosis, single dispatch, co-dispatch, faster vehicle, two vehicles collaboration, and fusion. In Table 1, we present the simulation results concerning the average time to diagnosis and time to treatment for the described scenarios and the current situation in the SHR, where no MSU is used. Please note that the transferring to hospital policy is only applicable to the co-dispatch policy.

As shown in Table 1, by placing one MSU in ambulance site ID {3} and using co-dispatching, the average time to diagnosis decreases compared to the single dispatch policy (either RA-dispatch or MSU-dispatch) for the patients receiving thrombolysis, hemorrhagic treatment, or no treatment. Also, by using co-dispatching, the average time to treatment decreases compared to the single dispatch policy (either RA-dispatch or MSU-dispatch) for ischemic stroke patients who undergo at least one treatment. The results in Table 1 also demonstrate that the co-dispatch policy is used for 9.34% of all operations, and in the MSU-dispatch, the share of MSU operations is 6.86%.

When three MSUs are placed in the ambulance sites {2, 3, 38}, by using the co-dispatch policy, the average time to diagnosis decreases compared to the single dispatch policy (either RA-dispatch or MSU-dispatch) for all stroke patients except those treated with thrombectomy. Also, by using co-dispatching, the average time to treatment decreases compared to the single dispatch policy (either RA-dispatch or MSU-dispatch) for all patients. The results indicate that the co-dispatch policy is used for 61.40% of all operations, and the MSU-dispatch for 52.40%.

The outcomes in Table 1 indicate that while placing only one MSU in the SHR has a positive impact on patients' conditions compared to the current situation, the most favorable results for the average time to treatment among patients undergoing different treatments are achieved when three MSUs are placed in the specified locations within SHR. As shown in Table 1, by using co-dispatching, the RA arrives before the MSU at the patient's location in both MSU settings, leading to a longer average waiting time for the RA compared to the MSU. Furthermore, both transferring to hospital policies, that is, faster vehicle and two vehicles collaboration, yield the same results.

Table 1. Comparison of the average time to diagnosis and treatment (in hours) for each stroke treatment for the considered dispatch policies and transferring to hospital policies. MSU loc.: MSU locations; Disp.: dispatch policy; SD: single dispatch, CD: co-dispatch, THS: thrombolysis, THY: thrombectomy, THSY: thrombolysis + thrombectomy, HT: hemorrhagic treatment, NT: no treatment, TrHP: transferring to hospital policy, TrF: faster vehicle, TrT: two_vehicles_collaboration, WEV1: waiting time for the first arrival EV in hours, WEV2: waiting time for the second arrival EV in hours, and SH: share of the co-dispatch/MSU operations out of all operations. The numbers within the curly brackets show the ambulance site IDs (see Figure 5).

MSU loc.	Disp.	TrHP	Average time to diagnosis (h)					Average time to treatment (h)					WEV1 (h)	WEV2 (h)	SH (%)
			THS	THY	THSY	HT	NT	THS	THY	THSY	HT	NT			
-	SD (RA)	-	0.97	1.07	1.06	0.98	0.99	1.56	2.19	2.11	1.57	1.57	-	-	-
{3}	SD (MSU)	-	0.96	1.14	1.10	0.97	0.97	1.53	1.96	2.01	1.59	1.58	-	-	6.86
{3}	CD	TrT/ TrF	0.95	1.13	1.09	0.96	0.96	1.51	1.86	1.99	1.58	1.58	0.18	0.07	9.34
{2,3,38}	SD (MSU)	-	0.78	1.28	1.10	0.78	0.78	1.23	1.84	1.75	1.43	1.44	-	-	52.4
{2,3,38}	CD	TrT/ TrF	0.70	1.27	1.05	0.71	0.72	1.14	1.66	1.70	1.36	1.38	0.15	0.1	61.4

The CT scanner in the MSU is employed to distinguish between ischemic and hemorrhagic stroke types. However, it does not specify the patient's specific need for thrombolysis and thrombectomy or thrombectomy alone. As a result, we assumed that ischemic stroke patients potentially requiring thrombectomy should be transported to a hospital for further diagnosis, even if they have been initially diagnosed inside the MSU. The time to diagnosis for these patients is from when the emergency center receives a call until the patient is delivered to the hospital (either the closest hospital or SC). This justifies the reason that the obtained time to diagnosis for these patients using the co-dispatch policy in both MSU settings is longer than in the single-dispatch policy. The other reason is that we assumed that an RA drives faster than an MSU, thus, an RA transports the patient faster to the chosen hospital.

The results in Table 1 demonstrate the considerable waiting times for the involved EVs in the co-dispatch policy, particularly for the RA, the EV that arrives earlier at the meeting point in our scenario. Our framework currently uses precalculated travel times, limiting us to consider the patient's location as the meeting point for the two involved EVs in the co-dispatch policy in prehospital stroke care. To minimize the waiting time, we plan to develop the framework by dynamically calculating travel times between any two locations within the region under study. This enables us to dynamically determine optimal meeting points based on the current locations of the patient, RA, and MSU, leading to a faster and more efficient service for the patient.

4. Conclusions

In this paper, we presented a simulation-based analysis to study the potential impacts of using a co-dispatch policy involving both RAs and MSUs in the prehospital transportation of stroke patients. We included the co-dispatch policy and different types of stroke to a modeling framework, introduced in our previous study [4]. The proposed framework is sufficiently general for application in any region, allowing end-users to adapt it to the region of interest by providing the appropriate inputs. In the scenario study, we applied the extended framework to southern Sweden to assess the benefits of using the co-dispatch policy for stroke patients. Our study, including different stroke types and scenarios, provided a comprehensive insight into how a co-dispatch policy might affect transporting stroke patients before they reach the hospital. The simulation results demonstrated the advantages of using the co-dispatch policy by indicating reductions in time to both diagnosis and treatment for stroke patients receiving different types of treatment.

In future work, we plan to incorporate real-time conditions such as traffic and weather phenomena into the model, enabling more realistic modeling of prehospital stroke transportation and care.

References

[1] World Stroke Organization, "Facts and figures about stroke," [Online]. Available: https://www.world-stroke.org/world-stroke-day-campaign/why-stroke-matters/learn-about-stroke/. [Accessed: Sep. 20, 2023].

[2] M. Ebinger, B. Winter, M. Wendt, J. E. Weber, C. Waldschmidt, M. Rozanski, A. Kunz, P. Koch, P. A. Kellner, and D. Gierhake, "Effect of the use of ambulance-based thrombolysis on time to thrombolysis in acute ischemic stroke: a randomized clinical trial," *Jama*, vol. 311, no. 16, pp. 1622-1631, 2014.

[3] S. Amouzad Mahdiraji, J. Holmgren, R.-C. Mihailescu, and J. Petersson, "A micro-level simulation model for analyzing the use of MSUs in Southern Sweden," *Procedia Computer Science*, vol. 198, pp. 132-139, 2022.

[4] S. Amouzad Mahdiraji, J. Holmgren, A. A. Alshaban, R.-C. Mihailescu, J. Petersson, and J. Al Fatah, "A framework for constructing discrete event simulation models for emergency medical service policy analysis," *Procedia Computer Science*, vol. 210, pp. 133-140, 2022.

[5] Alassadi, F. Lorig, and J. Holmgren, "An agent-based model for simulating travel patterns of stroke patients," *DIGITAL 2021 – Advances on Societal Digital Transformation*, pp. 11-16, 2021.

[6] S. Mathur, S. Walter, I. Q. Grunwald, S. A. Helwig, M. Lesmeister, and K. Fassbender, "Improving prehospital stroke services in rural and underserved settings with mobile stroke units," *Frontiers in Neurology*, vol. 10, 2019.

[7] The Swedish Stroke Register (Riksstroke), "Annual Report from Riksstroke: Stroke and TIA," [Online]. Available: https://www.riksstroke.org/sve/forskning-statistik-och-verksamhetsutveckling/rapporter/arsrapporter/. [Accessed: Sep. 15, 2023].

[8] S. Amouzad Mahdiraji, J. Holmgren, R. C. Mihailescu, and J. Petersson, "An optimization model for the tradeoff between efficiency and equity for mobile stroke unit placement," In *Innovation in Medicine and Healthcare: Proceedings of 9th KES International Conference on Innovation in Medicine and Healthcare (KES-InMed-21)*, pp. 183-193, 2021.

PAPER VII - IMPLEMENTING DYNAMIC TRAVEL TIME CALCULATION IN EMS SIMULATIONS: IMPACTS ON PREHOSPITAL STROKE CARE AND TRANSPORTATION

Saeid Amouzad Mahdiraji, Marcus Juninger, Nicholas Narvell, Johan Holmgren, Radu-Casian Mihailescu, and Jesper Petersson

ABSTRACT

Preparing travel time data can be a time-consuming process, which greatly limits the flexibility of transport simulation models. In the current paper, we present an approach to integrate a routing engine locally in an existing modeling framework, hence enabling to dynamically calculate travel times in the constructed emergency medical services (EMS) simulation models. This integration eliminates the need for the pre-calculation typically required to prepare travel time data. Using the extended framework, we developed an EMS simulation model for stroke patients, which we applied in a scenario study to southern Sweden. This allowed us to evaluate the potential benefits of using dynamic travel time calculations in prehospital stroke care. The experimental results, supported by comparisons with pre-calculated travel times, confirm the effectiveness of our approach in integrating dynamic travel time calculations into the framework. Moreover, the results of our evaluation indicate that including this functionality in simulation models can provide more realistic results. Finally,

our approach for local implementation of dynamic travel time calculations is faster and less restricted compared to using online services.

PUBLISHED IN

Accepted for Publication in Proceedings of the International Conference on Health and Social Care Information Systems and Technologies (HCist 2024).

1. Introduction

Transport simulation is an important tool for analyzing and planning transport systems, enabling, for example, the prediction of traffic patterns and the assessment of infrastructure projects. By simulating different transport scenarios, decision-makers can use simulation models to evaluate the impact of different policies and changes before real-world implementation. The dynamic nature of transport systems, characterized by changes in traffic conditions and unexpected road closures, necessitates the use of dynamic data to ensure the accuracy and relevance of simulation outcomes.

Before or during the simulation, different types of data, such as travel times, may need to be provided to the transport simulation model. Some methods for preparing travel times in transport simulation models involve calculating travel times as origin-destination matrices in advance for all potential locations within the study region. These pre-calculated matrices are then used in the simulation model during execution. However, such travel time data preparation approaches are complex and time-consuming since some of the data preparation tasks typically need to be performed manually. Moreover, due to the complex and long data preparation process, the travel data cannot be frequently updated and are instead used as static data. An important limitation of static travel times data is that they fail to account for real-time traffic conditions in the simulation models.

In response to these challenges, dynamic travel time calculations have been introduced to simulation models to provide more accurate and dynamic data that reflect the current transport conditions. The integration of dynamic travel time calculations into the transport simulation models allows the model to efficiently adapt and respond to variations in input, for example, resource availability and changes in traffic patterns, and to simplify the data preparation process. Another advantage is that, by using the dynamic approach, only the required travel data will be calculated. This integration would also improve the predictive accuracy of simulations [1]. In this regard, Hajinasab et al. [2] propose using online services to outsource the generation and calculation of data for transport simulation. This approach is advantageous as it eliminates the need to pre-calculate vast amounts of travel data. Huber and Rust [3] introduce an approach to locally calculate thousands of travel time requests within seconds. As these calculations are performed offline, their approach can effectively use the full computational power of the host system and support a large number of requests.

The need for dynamic travel data is crucial in emergency medical services (EMS) simulations, where the speed and efficiency of prehospital interventions can critically impact patient outcomes, for example, for acute conditions like stroke. EMS simulation models, used to optimize operations and improve patient care, may use pre-calculated, static, travel times. While these static times are useful for basic scenario planning, they fail to reflect the unpredictable and dynamic nature of real-world conditions that emergency vehicles (EVs) encounter. The integration of dynamic travel time calculations into EMS simulation models can be also positive, showing significant improvements in the accuracy of the models and their applicability for policy-making. Patel et al. [4] and Zhen et al. [5] demonstrate the limitations of using Geographic Information System and other static data sources for modeling pre-hospital and ambulance deployment times, which often lead to suboptimal resource positioning during peak traffic periods or other dynamic events. Juninger and Narvell [6] evaluate routing engines for dynamic travel time calculations used in the simulation of EV transportation. Juninger and Narvell demonstrate that the most suitable routing engine is the *Open Source Routing Machine with Contraction Hierarchies* (OSRM-CH)[1]. They also realize that locally deploying the instances of the routing engines leads to better performance than running them online.

Some studies use pre-calculated travel data to simulate prehospital stroke care and transportation [7, 8]. However, the impact of including dynamic travel times in EMS simulations, particularly for prehospital stroke care, remains largely unexplored. In this paper, we build on the approach by Juninger and Narvell [6] for dynamic travel time calculations. In particular, we integrate the OSRM-CH routing engine into an established EMS modeling framework, used to construct EMS simulation models for different medical conditions, including stroke [7]. This integration enhances the framework's capability to calculate travel times dynamically, removes the need to pre-calculate travel data, and it aims to better align the simulation outcomes with actual EMS operations. We explore the benefits and limitations of our proposed approach through a scenario study applied to southern Sweden.

The rest of the paper is structured as follows. In Section 2, we present how we integrate dynamic travel time calculations into the modeling framework and then use them to construct a stroke simulation model [7]. Section 3 describes the scenario study, followed by the results and a discussion. Eventually, Section 4 concludes the paper.

[1] https://github.com/Project-OSRM/osrm-backend/

2. Integrating Dynamic Travel Time Calculations in a Modeling Framework

In this section, we present our framework for building EMS simulation models, which is followed by a description of the integration of the dynamically calculated travel times into the framework.

2.1 Modeling Framework

In a prior study [7], we introduced a modeling framework designed to construct EMS simulation models. As illustrated in Figure 1, the framework comprises two phases: model construction and model execution. In the construction phase, the framework uses different inputs, including geographical data, patient data, EV data, hospital data, and a care chain specification, to build a discrete event simulation model tailored to a specific EMS scenario. During the execution phase, the constructed model is executed using the input data, and the specified care chain is applied to each simulated patient. In the original version of the framework, pre-calculated travel times of EVs are given as a static input to the framework. Considering the potential advantages of using dynamic travel time calculations during simulation, we below describe how we incorporate this functionality into the framework.

There are five different types of transport activities in the framework that require travel time calculations: 1) travel from the current location of an EV to the patient's location, 2) travel from the patient's location to a hospital, 3) travel

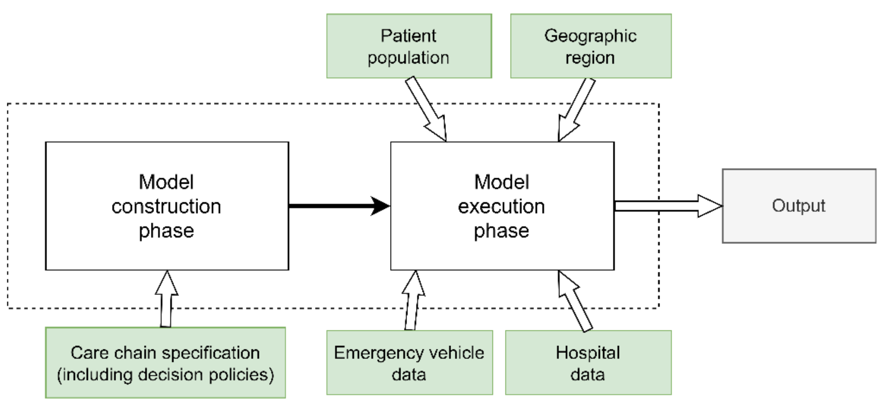


Figure 1. Overview of the modeling framework proposed by Amouzad Mahdiraji et al. [7], where the provided input data is used to construct and to run the model.

from a hospital to an EV site, 4) travel from the patient's location to an EV site, and 5) travel from a hospital to a special clinic. Each transport activity type requires a unique way of calculating travel times. Transport activity type 1 involves an initial step to determine which EV(s) should be dispatched to the patient's location. This decision-making process requires calculating the travel times from all EV sites to identify which EV is expected to arrive fastest at the patient location. However, performing such calculations from the current locations of all EVs to the patient's location using pre-calculated ambulance driving times would be complicated. Transport activity type 2 necessitates calculating travel times from the patient's location to all hospitals or special clinics to determine the nearest one. These two activities require the calculations of many-to-one and one-to-many trips, respectively. Transport activity types 3, 4, and 5 only require one-to-one travel time calculations. For example, for transport activity types 3 and 4, each EV is typically stationed at and returns to the same EV site. Also, for transport activity type 5, there is generally only one special clinic in each region. However, there could be different special clinics for different diagnoses and treatments.

2.2 Integrating Dynamic Travel Time Calculations into the Framework

In the original version of the framework [7], we pre-calculated the travel times of EVs using the driving time generation feature of OpenStreetMap (OSM), which we accessed through the Openrouteservice toolbox in QGIS. OSM, created by the OpenStreetMap Foundation, is based on geographic data and provides a free editable world map [9]. Due to the limited number of requests allowed by OSM per run, we needed to manually divide the input data, including EV data, hospital data, and patient data, before processing it through the Openrouteservice toolbox. Subsequently, we merged all the calculated travel data in QGIS to compile the travel time data set. This method of calculating travel times proved to be inconvenient, time-consuming, and prone to errors, particularly when adapting the framework to new regions. In addition, using the pre-calculated travel data limits the framework in several ways. For example, it requires that EVs are always dispatched from their sites to the patients' locations. Hence, we could not account for the possibility of dispatching an EV from any location within the region. It also requires that an EV must return to its site from the hospital or patient's location to be ready for the next call. A dynamic approach for travel data calculation can overcome these limitations by allowing for more flexible and accurate travel time calculations. Due to the large number of travel time calculations required by the framework, the chosen approach should not restrict the number of travel times to be calculated.

To enable the framework to calculate the EV travel time dynamically during simulation, we include the approach by Juninger and Narvell [6] into the framework. Juninger and Narvell identify and evaluate a set of routing engines based on several criteria, such as license price, request limitations, matrix search ability, performance, and usability. Among the available routine engines, the authors identify that ORS, OpenSourceRoutingMachine (OSRM), and Valhalla offer unlimited requests for local deployment. For evaluation, the authors set up two so-called Docker instances for each routing engine: one hosted locally and one hosted at a separate location. Docker[2] is a platform that uses lightweight virtualization to package applications into containers, allowing them to run consistently across different computing environments. This containerization simplifies deployment and supports scalability by enabling multiple instances of the same application to operate simultaneously. They also develop a performance test suite to log statistics to calculate and return requests for each routine engine. The experimental results demonstrate that OSRM-CH outperforms other routing engines for local deployment. OSRM-CH is a C++-based routing engine that uses Contraction Hierarchies as its traversal algorithm. Furthermore, they integrate the OSRM-CH into a module[3] that makes it possible to dynamically calculate travel times.

In this paper, we utilize the module by Juninger and Narvell to locally integrate the OSRM-CH into the framework. In Figure 2, we illustrate how the framework and the module interact to calculate the travel times during the simulation. The OSM data for the region of interest is downloaded from Geofabrik[4]. To enable the module to function, an instance of OSRM_CH, which is preloaded with OSM data, is configured and run locally in Docker. The module restructures the input data to align with the routing engine's interface before sending HTTP requests to the Docker container where the routing engine is running.

The patient data, hospital data, and EV data, fed into the framework as input data, should include either a single coordinate (a pair of longitude and latitude) or a list of coordinates. The module handles three types of trip requests based on the transport activity types, as described in Table 1. Each request type has two input parameters: *origin* and *destination*, and the return value is a tuple containing

[2] https://www.docker.com/
[3] https://github.com/Juninger/python-travel-time-module/
[4] https://download.geofabrik.de/

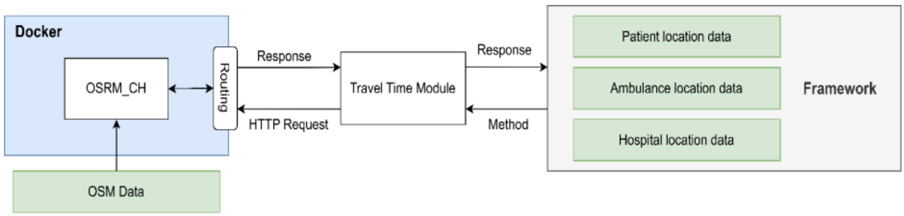


Figure 2. An overview of the data flow and relationships between the framework and the module [6] for the dynamic travel time calculations.

the fastest travel time and the corresponding coordinates for the origin and destination, that is, *{travel_time, [long, lat], [long, lat]}*. Considering the transport activity types included in the framework, transport activity type 1 is modeled using the many-to-one request type, transport activity type 2 using the one-to-many request type, and transport activity types 3-5 using the one-to-one request type.

The proposed approach can dynamically calculate travel times for different scenarios applicable to the framework. All these calculations can occur during the framework's runtime based on the requested trip type, eliminating the need for any pre-calculated travel times.

Table 1. Description of each type of trip request available in the module by Juninger and Narvell [6].

Request type	Description
Many-to-one	The origin parameter takes a list of coordinates, that is, *origins: [[long, lat], [long, lat], ...]*. The destination parameter takes a list of a single coordinate, that is, *destination: [long, lat]*.
One-to-many	The origin parameter takes a list of a single coordinate, that is, *origin: [long, lat]*. The destinations parameter takes a list of coordinates, that is, *destinations: [[long, lat], [long, lat], ...]*.
One-to-one	The origin parameter takes a list of a single coordinate, that is, *origin: [long, lat]*. The destination parameter takes a list of a single coordinate, that is, *destination: [long, lat]*.

3. Scenario Study

We evaluated the extended framework by constructing a simulation model for prehospital stroke care, which we applied to Sweden's southern healthcare region (SHR). In the current section, we provide a brief introduction to stroke and describe the scenario study, as well as the framework configurations, before presenting the experiments.

Stroke is known as one of the main causes of death and disability worldwide. Ischemic stroke is the most common type of stroke, where thrombolysis is the standard treatment. Timely treatment is essential for the successful recovery of stroke patients; however, logistical challenges often delay the delivery of appropriate treatment. In addition, stroke patients are typically transported to an acute hospital for diagnosis and treatment, which can be time-consuming. However, earlier studies demonstrate that the use of mobile stroke units (MSUs) has the potential to significantly reduce the time to treatment for stroke patients. MSUs are specialized EVs equipped with a CT scanner that enables both the diagnosis and administration of thrombolysis directly on-site [10]. An MSU may be deployed either alone or alongside a regular ambulance (RA) in response to a stroke incident.

In the scenario study, we applied the proposed approach of integrating dynamic travel time calculations into the framework to SHR, illustrated in Figure 3. As of 2023, the total population of SHR was 1,926,100 [11]. According to the Swedish Stroke Register (Riksstroke) [12], Sweden recorded over 20,000 annual stroke incidents in 2022, including 3,900 cases within SHR. Figure 3 illustrates the locations of ambulance sites and acute hospitals in the SHR, marked by green triangles and purple circles, respectively, with each ambulance site assigned a unique ID. Currently, SHR does not have any MSUs in operation.

We conducted several simulation runs using the constructed simulation model for prehospital stroke care on a computer equipped with 32 gigabytes of RAM and an Intel(R) Core(TM) i7-8650U CPU operating at 1.90 gigahertz. The framework settings allow users to run experiments with or without using the dynamic travel time calculation functionality. In the simulation runs, we considered decision policies concerning: 1) the selection of which EV to be dispatched to the patient's location, 2) the choice between single dispatch and co-dispatch, and 3) the selection of the destination hospital. For EV selection, we adopted the *time to diagnosis* policy, where one or two EVs—referred to as single dispatch and co-dispatch, respectively—are chosen to minimize the expected time until diagnosis. For hospital selection, our choice was the *fusion* policy, allowing the patient to be transported either to the nearest hospital or to a special clinic based on the on-scene assessment (and diagnosis) and the distance from the patient's location to the special clinic. Regarding the dispatch type, the user can choose between *single dispatch* and *co-dispatch*. The single dispatch policy refers to the situation where either an RA or an MSU is dispatched to a stroke incident. Co-dispatching, on the other hand, refers to when an RA and an MSU collaborate until the suspected stroke patient receives appropriate treatment.

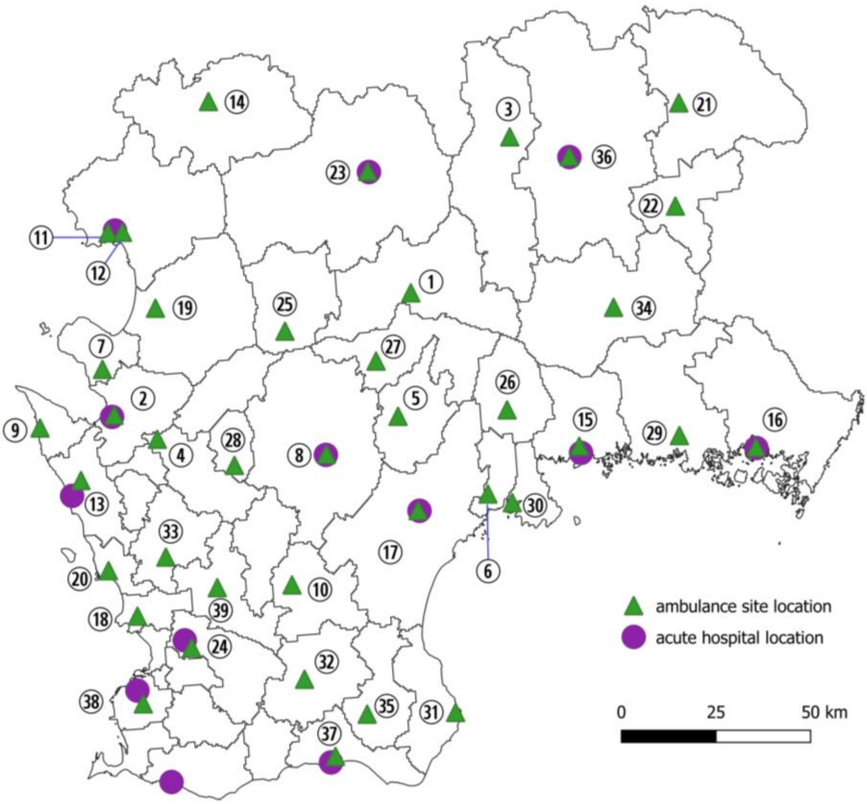


Figure 3. Overview of Sweden's southern healthcare region, reproduced from [15]. The purple circles and green triangles represent the locations of acute hospitals and ambulance sites, respectively. The circled numbers indicate the corresponding ambulance site IDs.

We utilized the same input data, MSU locations, and assumptions as in our previous study [13]. The geographic region, that is, the SHR, is divided into non-overlapping 1×1 km^2 subregions to facilitate the localization of patients, hospitals, and ambulance sites. The stroke patient data, generated using a Poisson distribution, represents a synthetic population distributed across the SHR (details are provided in a previous study [14]). Each patient in this population is characterized by specific attributes including incident time, location (coordinate), age, symptoms, diagnoses, and preferred treatment. We assumed that all patients represented in the simulation are ischemic stroke patients requiring diagnosis and, potentially, treatment either inside the MSU or at a hospital. Regardless of the diagnosis and treatment location, the patients would ultimately be transported to a hospital for further care. In all cases, the time to diagnosis or treatment is defined as the expected duration from the onset of a stroke until the patient receives diagnosis or treatment, either inside the MSU or at a hospital.

We took into account two scenarios that correspond to the dispatch policies: 1) single dispatch using RA only, which reflects the current situation in the SHR, and 2) co-dispatch. We evaluated these scenarios in relation to other decision policies, specifically ambulance selection and hospital selection, and under both dynamic and pre-calculated travel time conditions. In our experiments, the locations of the MSUs in the SHR were based on a previous study [15], in which we proposed an optimization model for the MSU placement problem to make a tradeoff between assuring equitable service and maximizing population coverage. We therefore set up one MSU in Alvesta (ID: 3), two MSUs in Malmö (ID: 38) and Alvesta (ID: 3), or three MSUs in Ängelholm (ID: 2), Alvesta (ID: 3), and Malmö (ID: 38) to conduct the experiments.

3.1 Experimental Results and Discussion

In Table 2, we present the simulation results for the constructed model with and without dynamic travel time calculations concerning the average time to diagnosis and treatment for the described scenarios and policies. The results show that the use of co-dispatching, compared to the RA-dispatch, that is, the current situation in the SHR, is expected to reduce the time to diagnosis and treatment for stroke patients in the SHR, particularly when three MSUs are involved in co-dispatch operations, the expected reduction is 15.6 and 24.6 minutes for time to diagnosis and treatment, respectively.

According to Table 2, the time to diagnosis and treatment are approximately equal for dynamic and pre-calculated travel times, demonstrating that the proposed approach for integrating the dynamic travel time calculations of EVs into the framework functions as intended. However, using the dynamic travel time data results in a slightly longer time to diagnosis and treatment for all scenarios. The slight differences in the results arise because the dynamic travel times calculated using OSRM_CH reflect the most recent changes in road networks. In addition, the pre-calculated travel data was created using a different tool, namely the Openrouteservice toolbox in QGIS.

Furthermore, in Table 2, we compare the execution time of the simulation model applied to the SHR both with and without dynamic travel time calculations. The comparison shows that the running time using dynamic travel time calculations is higher since the framework, for calculating each travel

Table 2. Comparison of running time (in hours), loading time for pre-calculated data (in hours), and the average time to diagnosis and treatment (in hours) for the considered dispatch policies both with and without dynamic travel times. The numbers within the curly brackets show the ambulance site IDs (see Figure 3). SD: single dispatch, CD: co-dispatch, ATD: average time to diagnosis, and ATT: average time to treatment.

MSU locations	Dispatch policy	Travel time calculation	ATD (h)	ATT (h)	Running time (h)	Loading time of pre-calculated data (h)
-	SD (RA)	Static	1.01	1.59	0.44	0.37
-	SD (RA)	Dynamic	1.02	1.60	1.05	-
{3}	CD	Static	0.98	1.54	0.43	0.36
{3}	CD	Dynamic	0.99	1.55	1.13	-
{3, 38}	CD	Static	0.83	1.29	0.45	0.36
{3, 38}	CD	Dynamic	0.84	1.31	1.15	-
{2, 3, 38}	CD	Static	0.74	1.18	0.46	0.36
{2, 3, 38}	CD	Dynamic	0.76	1.19	1.16	-

activity, interacts with the module, which sends the requests to the Docker container. This process would take a few milliseconds for each request. Conversely, the time spent to calculate travel times data in advance for simulation is enormous, especially when the region under study is large, restricting the model's flexibility to consider different simulation scenarios. In addition, loading the pre-calculated travel data of SHR to the framework before starting the simulation would take about 22 minutes on average, which also needs to be taken into consideration. In this condition, calculating travel times data for a larger region and loading it into the model would be even more time-consuming. Hence, the use of pre-calculated data for real-world EMS situations is not beneficial. As a result, while adding dynamic travel time calculation functionality to the framework slightly increases the running time, it eliminates the need to prepare travel data for the simulation, which is a very time-consuming task.

To further evaluate our approach for dynamically calculating travel times in an EMS simulation model for stroke patients, we compare the running time of our approach, which is based on a local routing engine, with that of an online service, that is, the ORS Web API[5]. ORS Web API provides a free online routing service with a limited number of requests per minute and day, that is, 500 matrix route calculations. The volume of requests does not support our scenario study. However, to make a comparison between the performance of ORS Web API and our proposed approach, we randomly chose 500 pairs of coordinates between patient locations and ambulance locations in the SHR. Then, we divided these

[5] https://openrouteservice.org/plans/

500 requests into 13 batches, where each batch of 40 requests is sent to ORS Web API per minute to calculate the travel times. The experiment reveals that calculating 13 batches of a total of 500 requests sent to the ORS Web API took 316.60 seconds. In contrast, executing the same number of requests using the local routing engine used in this paper required 5.11 seconds. Consequently, the online service is 44 times slower than the local routing engine integrated into the framework.

4. Conclusions

We have explored the potential advantages of dynamic calculation of travel times in EMS simulation models. For this aim, we used the proposed approach by Juninger and Narvell [6] for dynamic travel time calculations and integrated it into a modeling framework for construction of EMS simulation models. The chosen routing engine was OSRM_CH. In our scenario study, we created an EMS simulation model for the stroke care chain, which we applied to southern Sweden, and compared the effectiveness of our approach with or without dynamic travel data.

The experimental results, supported by the results using pre-calculated travel data, showed that the proposed approach for dynamic travel data calculations works successfully in the EMS simulation model for stroke patients. The simulation output analysis demonstrated that although the use of pre-calculated travel data can lead to shorter execution times, the time to calculate travel data and to load it during simulation would be considerable. However, when using dynamic travel data calculations during simulation, we do not need to calculate the travel times in advance, which is a time-consuming task, especially when the region of study is vast. Also, it is more straightforward to adapt the framework to new regions only by providing the required input data, including ambulance locations, patient locations, and hospital locations, and a care chain specification. Comparing the local-based and online services for dynamic travel time calculations showed that the local implementation, used in our approach, is faster and less restricted regarding the number of requests compared to online services.

References

[1] R. Fujimoto, J. Barjis, E. Blasch, W. Cai, D. Jin, S. Lee, and Y.-J. Son, "Dynamic data driven application systems: research challenges and opportunities," in *2018 Winter Simulation Conference (WSC)*, pp. 664-678, IEEE, 2018.

[2] B. Hajinasab, P. Davidsson, J. Holmgren, and J. A. Persson, “On the use of online services in transport simulation,” *Transportation Research Procedia*, vol. 21, pp. 208-215, 2017.

[3] S. Huber and C. Rust, “Calculate travel time and distance with OpenStreetMap data using the Open Source Routing Machine (OSRM),” *The Stata Journal*, vol. 16, no. 2, pp. 416-423, 2016.

[4] B. Patel, N. M. Waters, I. E. Blanchard, C. J. Doig, and W. A. Ghali, “A validation of ground ambulance pre-hospital times modeled using geographic information systems,” *International Journal of Health Geographics*, vol. 11, pp. 1-10, 2012.

[5] L. Zhen, K. Wang, H. Hu, and D. Chang, “A simulation optimization framework for ambulance deployment and relocation problems,” *Computers & Industrial Engineering*, vol. 72, pp. 12-23, 2014.

[6] M. Juninger and N. Narvell, “On the use of routing engines for dynamic travel time calculation within emergency vehicle transport simulation,” *bachelor thesis,* Malmö University, 2023.

[7] S. Amouzad Mahdiraji, J. Holmgren, A. A. Alshaban, R.-C. Mihailescu, J. Petersson, and J. Al Fatah, “A framework for constructing discrete event simulation models for emergency medical service policy analysis,” *Procedia Computer Science*, vol. 210, pp. 133-140, 2022.

[8] J. Al Fatah, A. A. Alshaban, J. Holmgren, and J. Petersson, “An agent-based simulation model for assessment of prehospital triage policies concerning destination of stroke patients,” *Procedia Computer Science*, vol. 141, pp. 405-412, 2018.

[9] Open Street Map, “About OpenStreetMap,” [Online]. Available: https://wiki.openstreetmap.org/wiki/About_OpenStreetMap. [Accessed: April 1, 2024].

[10] M. Ebinger, B. Winter, M. Wendt, J. E. Weber, C. Waldschmidt, M. Rozanski, A. Kunz, P. Koch, P. A. Kellner, and D. Gierhake, “Effect of the use of ambulance-based thrombolysis on time to thrombolysis in acute ischemic stroke: a randomized clinical trial,” *Jama*, vol. 311, no. 16, pp. 1622-1631, 2014.

[11] Statistics Sweden, “Demographic data 2023,” [Online]. Available: https://www.scb.se/. [Accessed: April 1, 2024].

[12] The Swedish Stroke Register (Riksstroke), “Annual Report from Riksstroke: Stroke and TIA,” [Online]. Available: https://www.riksstroke.org/sve/forskning-statistik-och-verksamhetsutveckling/rapporter/arsrapporter/. [Accessed: Sep. 15, 2023].

[13] S. Amouzad Mahdiraji, J. Holmgren, R. C. Mihailescu, and J. Petersson, “Simulation-based analysis of co-dispatching in prehospital stroke care,” *Procedia Computer Science*, vol. 238, pp. 412-419, 2024.

[14] S. Amouzad Mahdiraji, J. Holmgren, R.-C. Mihailescu, and J. Petersson, “A micro-level simulation model for analyzing the use of MSUs in Southern Sweden,” *Procedia Computer Science*, vol. 198, pp. 132-139, 2022.

[15] S. Amouzad Mahdiraji, J. Holmgren, R. C. Mihailescu, and J. Petersson, “An optimization model for the tradeoff between efficiency and equity for mobile stroke unit placement,” In *Innovation in Medicine and Healthcare: Proceedings of 9th KES International Conference on Innovation in Medicine and Healthcare (KES-InMed-21),* pp. 183-193, 2021.

PAPER VIII - INTEGRATING MACHINE LEARNING-BASED AMBULANCE TRAVEL TIME ESTIMATION INTO AN EMERGENCY MEDICAL SERVICES SIMULATION MODELING FRAMEWORK

Saeid Amouzad Mahdiraji, Muhammad Adil Abid, and Johan Holmgren

ABSTRACT

Travel time estimation is an integral component of emergency medical services (EMS) simulations due to the need to calculate ambulance transport times for patients. We present a study where we integrated a machine learning (ML) based ambulance travel time estimation module into an EMS simulation modeling framework, aiming to explore the potential benefits of using ML-based travel time estimations in emergency simulations. To illustrate the effectiveness of the proposed approach, we used the framework to construct an EMS simulation model for stroke patients and applied it in a scenario study covering Skåne County, Sweden. The result of the simulation shows differences in ambulance driving times when using the ML-based module compared to existing routing engines designed for passenger cars. The observed differences emphasize the impacts of integrating ML-based estimations into EMS simulations.

PUBLISHED IN

Proceedings of the 14th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH 2024), vol. 251, pp. 479-486, 2024.

1. Introduction

Emergency medical services (EMS) focus on the prehospital stabilization, treatment, and transportation of acutely ill or injured patients. Simulation models are increasingly employed to enhance decision-making and operational efficiency in EMS, and they offer a cost-effective way of evaluating and optimizing various aspects of EMS operations, including training, resource allocation, and strategic planning [1].

The accurate estimation of ambulance travel times between various locations within a geographic region is crucial for the reliability of EMS simulations. Ambulance travel time estimation predicts the time required for an ambulance to travel between two locations, for example, from an emergency scene to a hospital. Factors like traffic conditions, time of day, day of the week, weather, and the characteristics of the road network can greatly affect the travel time. Failing to consider these factors can result in inaccurate ambulance travel time estimations. For example, the geographically closest ambulance might not be the fastest to reach a patient due to the current traffic or road conditions. Furthermore, in time-sensitive emergencies like stroke, selecting the best ambulance and hospital based on accurate travel time estimations is vital.

We argue that incorporating more realistic travel time estimations in EMS simulations has the potential to enhance decision-making, leading to better resource allocation, quicker medical responses, and improved patient outcomes. Accurate travel time estimations also make the simulations more reliable and effective. Accurately estimating travel time is a key challenge in simulation modeling due to the variability in traffic and road conditions [5]. Previous research suggests using routing engines like Google Maps and Open Source Routing Machine (OSRM) to estimate travel times during simulations [2-4]. However, these estimators are primarily designed for passenger cars and do not accurately estimate the travel time of ambulances. This mismatch can lead to inaccurate estimations, negatively impacting emergency response planning, resource allocation, and patient care. Therefore, it would be beneficial to use a specific travel time estimation model for ambulances in simulations that accounts for the unique travel patterns of ambulances and external conditions.

Recent advances in ML offer a promising alternative to conventional methods like static speed models, historical averages, and deterministic algorithms such as Dijkstra’s for shortest path calculations. While traditional optimization models are useful for route selection and shortest-path problems —for example, finding the optimal path in a transport network— ML approaches can be used to determine the actual driving time along a path or to correct the calculated travel times by the mentioned routing engines.

Several studies demonstrate the effectiveness of ML in enhancing ambulance travel time estimations. Torres et al. [6] propose a supervised ML model to enhance ambulance travel time predictions, which is applied to the Red Cross of Tijuana, Mexico. Torres et al. use their model to correct the calculated travel times provided by Google Maps and the OSRM. The approach involves compiling a dataset from EMS logs and global positioning system (GPS) data, then training a random forest classifier under three conditions: default settings, hyperparameter optimization, and using an AutoML system. Buna et al. [7] contribute an ambulance travel time estimation model for the Žilina Region of Slovakia, by analyzing GPS data from seventeen ambulances over three years. Their approach includes three steps: extracting raw GPS data, preprocessing the data to identify individual ambulance trips, and applying statistical models to capture the spatiotemporal characteristics of the identified trips. Buna et al. combine linear and nonlinear regression models, histograms, and probability distribution functions to create synthetic ambulance trips. Abid et al. [8] present an ML-based approach to estimate ambulance travel times using real-world spatiotemporal data from Skåne County, Sweden. Their approach includes explanatory data analysis, data preprocessing, feature selection, and ML modeling. To evaluate the effectiveness of their proposed approach, the authors analyze the performance of different ML models—polynomial regression, XGBoost, and artificial neural networks (ANN)—across various feature sets. Hu et al. [9] introduce a travel time estimation model for traffic corridors utilizing simulation-assignment methods. The authors propose two algorithms in this regard: a flow-based model that uses instantaneous link speeds and a vehicle-based model that uses simulated vehicle trajectories. The model integrates dynamic origin-destination data generated through a Kalman-filter-based approach to simulate traffic conditions within the DynaTAIWAN framework.

Several studies in EMS utilize the abovementioned routing engines, which provide passenger car travel data, during simulations to analyze prehospital care and transportation [2, 4, 10]. To the best of our knowledge, the benefits of using an ML-based approach to estimate ambulance travel times in simulations remain unexplored. In the current paper, we propose a new approach that integrates an ML-based ambulance travel time estimation module into an established simulation modeling framework, offering a comprehensive tool for EMS policy analysis and decision-making. In particular, we integrated the ambulance travel time estimation module proposed by Abid et al. [8] into an EMS modeling framework, used to build EMS simulation models for different medical conditions [10]. This integration is expected to improve the accuracy of the simulation outcomes, bringing them closer to real-world EMS operations and decision-making scenarios. We evaluate the proposed approach through a series of experiments conducted in a scenario study covering Skåne County, Sweden.

The remainder of this paper is organized as follows. In Section 2, we describe how we integrated an ML-based ambulance travel time estimation module into a simulation modeling framework, which is then used to build a stroke simulation model. In Section 3, we present our scenario study, experimental results, and a discussion. Finally, we conclude the paper in Section 4.

2. Integrating ML-Based Ambulance Travel Time Estimation in a Simulation Modeling

In this section, we first outline our framework for constructing EMS simulation models. We then describe the extended framework, in which we integrated an ML-based ambulance travel time estimation module.

2.1. Framework for EMS simulation model construction

In a prior study [10], we present a modeling framework for creating EMS simulation models. As shown in Figure 1, the framework consists of two main phases: model construction and model execution. During the construction phase, various inputs, such as geographical data, patient data, ambulance data, hospital data, and a care chain specification, are utilized to create a discrete event simulation model for a specific EMS scenario. In the execution phase, the constructed model is run with the input data, and the specified care chain is applied to each simulated patient.

The version of the framework proposed by Amouzad Mahdiraji et al. [2], which is an extension of [10], supports dynamic travel time calculations for ambulances and dispatch types policies (single dispatch or co-dispatch). It uses the OSRM with contraction hierarchies (OSRM-CH) routing engine for dynamic ambulance travel time calculations, eliminating the need for pre-calculated travel times, which was the case in the earlier versions of the framework.

The framework supports five transport activity types that require ambulance travel time calculations, and which are considered in the framework: 1) travel from the current location of an ambulance to the patient's location, 2) travel from the patient's location to a hospital, 3) travel from a hospital to an ambulance site,

4) travel from the patient's location to an ambulance site, and 5) travel from a hospital to another hospital; for example, to a special clinic. A dynamic travel time calculation module is used to calculate the travel time for each transport activity type. It should be noted that an ambulance site refers to a location where ambulances are stationed when they are not in transit or assigned to an operation.

Each patient in the population is specified using features consisting of incident time (hour, day of the week, day of the year), location (coordinate), age, sex, symptoms, priority level, diagnosis, and preferred treatment. The hospital and ambulance data includes hospital and ambulance site locations (coordinates) in the study region.

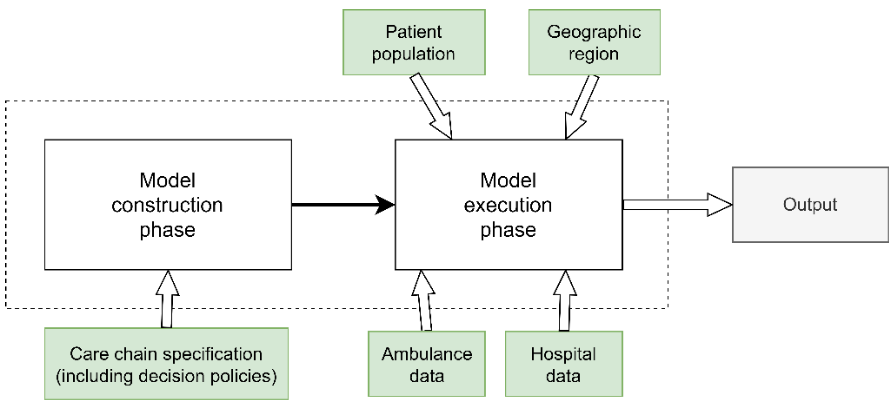


Figure 1. Overview of the modeling framework proposed by Amouzad Mahdiraji et al. [10], where the provided input data is used to construct and to run the generated simulation model.

2.2 Integration of ambulance travel time estimation module into the framework

As mentioned above, the version of the framework proposed by Amouzad Mahdiraji et al. [2] supports dynamic travel time calculations using the OSRM-CH routing engine. While integrating dynamic travel time calculations improves the accuracy and flexibility of EMS simulation models, it does not account for the accurate estimation of ambulance travel times, which is crucial for effective EMS simulations. To address this, a fixed percentage reduction is applied to the OSRM-CH travel times, assuming that ambulances drive faster in emergency situations.

The travel times calculated using OSRM-CH represent the average travel time for a passenger car, which likely differs from that of an ambulance. Due to the lack of data for all patient transports, we relied on generated travel times for cars to estimate ambulance travel times over the study region. Hence, we expect the dynamic travel time calculation adopted by Amouzad Mahdiraji et al. [2] might not fully reflect actual ambulance travel times in the simulation. Furthermore, the dynamic travel time calculation approach is not able to consider real-time factors such as traffic congestion, weather changes, road construction, and accidents, which can significantly affect ambulance speed and make the simulation less reliable. Additionally, since the calculation is based on passenger car travel times, it overlooks the fact that ambulances may exceed speed limits when safety allows, especially in emergencies where every second matters. Due to the limitations built into the dynamic travel time approach, there is great potential to improve EMS simulations by developing more accurate travel time estimations that better capture the speed and efficiency of emergency responses. This would lead to more reliable simulations, providing better decision support and ultimately improving patient outcomes.

In the current paper, we present a study where we integrated the ambulance travel time estimation module by Abid et al. [8] into the EMS simulation modeling framework by Amouzad Mahdiraji et al. [2] to explore its impacts on ambulance travel time estimations during simulations. Abid et al. [8] propose an ML-based approach for estimating ambulance travel times using spatiotemporal data. Their approach includes two phases: exploratory data analysis and a module for estimating ambulance travel times. The approach makes use of spatiotemporal data, divided into three categories: 1) ambulance spatial data, 2) temporal data, and 3) distance and calculated travel time data. The data analysis is performed to identify key features, patterns, and trends in the data that influence the ambulance travel times. The generation of the ambulance travel time estimation module consists of three main steps: preprocessing, feature selection, and ML modeling. In the ML modeling step, the authors compare the performance of various ML models across different travel scenarios and feature sets to enhance the accuracy of travel time estimations.

In this study, we compared the performance of three ML models—polynomial regression, XGBoost, and ANN—across various trip scenarios and feature sets. Based on the analysis, we selected the ANN as the best-performing ML model, using five features: 1) priority level from the ambulance spatial data, 2) hour of the day and 3) day of the week as temporal features, and 4) distance and 5) calculated travel time between geographic locations using OSRM-CH. We then integrated the best-performing ML model for ambulance travel time estimation into the simulation modeling framework. As illustrated in Figure 2, the ML model is used to convert the travel times calculated by the dynamic travel time calculation module into ambulance travel times during the simulation.

As presented in Figure 2, the patient data, hospital data, and ambulance data are fed into the framework. The integration process includes two phases. In phase one, the coordinates of the origin and destination for an ambulance transport activity are sent as a request to the dynamic travel time calculation module. The return value will be a tuple containing the fastest travel time, distance, and the coordinates for the origin and destination. In phase two, the travel time and distance from the returned tuple, hour of the day, day of the week, and priority level are sent as a new request to the ML-based ambulance travel time estimation module, which in turn returns the estimated ambulance travel time for the given request.

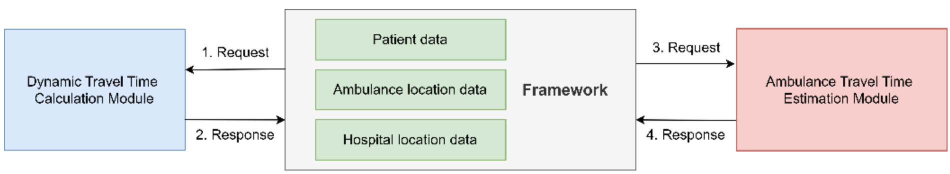


Figure 2. Illustration of how the framework and the ambulance travel time estimation module interact for estimating ambulance travel time.

3. Scenario Study

3.1 Study region description

To illustrate the proposed approach for integrating an ambulance travel time estimation module into our simulation modeling framework, we constructed a simulation model for stroke patients and applied it to Skåne County, Sweden. The Skåne County, which is shown in Figure 3, comprises 33 municipalities with a population of around 1.4 million in 2023 [11] and an area of 11,027 square kilometers. The region includes 8 acute hospitals and 22 ambulance sites.

Figure 3. Overview of the Skåne County in Sweden, where ambulance sites and hospitals are marked by red crosses and blue circles, respectively.

3.2 Data description and limitations

In this study, we retrained the ML model proposed by Abid et al. [8] using the same real-world spatiotemporal ambulance data. However, instead of using the Google Maps API for distance and travel time calculations, as done in the original study, we replaced it with OSRM-CH. This change aligns with our simulation modeling framework, which uses the OSRM routing engine to calculate passenger car travel times between two locations.

The ambulance spatiotemporal data (SOS Alarm), provided by the Skåne County of Sweden [12], encompasses the ambulance events from 2020 and 2021, describing ambulance transport activities corresponding to either emergency callouts or planned patient relocations. The data also includes pertinent information regarding the times for each ambulance trip, such as the travel times from the ambulance location to the patient and from the patient to the hospital.

Each ambulance event is assigned a priority level ranging from one (highest priority) to seven (lowest priority). In our analysis, we focused on the three highest priority levels, as the other levels represented a very small portion of the data. For further details, we refer the reader to [8].

To protect patient confidentiality, patient locations were reported by municipality names instead of exact coordinates. This added uncertainty to the data, particularly in larger municipalities, and could affect travel time estimations. There are also specific limitations related to transport activity types in the ambulance spatiotemporal data: i) for patient-to-hospital trips (transport activity type 2), the exact origin location was unknown and was given by a municipality name instead, while the exact destination was given; ii) for hospital-to-hospital trips (transport activity type 5), both the exact origin and destination locations were provided.

Abid et al. [8] compared the performance of ML models using various feature sets, where the feature set containing seven features (origin location, destination location, priority, hour of the day, day of the week, calculated distance, and calculated travel time) performed the best. However, in the current paper, we opted for the ML model with five features: priority, hour of the day, day of the week, distance, and calculated travel time. We excluded the origin and destination locations due to data limitations, including unknown exact coordinates for many trips and a limited number of travel records across the study region, which made these features less reliable. Additionally, in our simulation, the ML model often encounters many new locations that were not seen during training. Including origin and destination could bias the model toward the limited locations in the training data, reducing its generalizability. Furthermore, the start and destination information are implicitly captured through the distance and calculated travel time features, making it unnecessary to include them directly.

When applying the simulation modeling framework, we used the same input data and assumptions as in our previous study [2]. The stroke patient data consists of a synthetic population of approximately 3,000 patients, generated using a Poisson distribution and which are distributed across Skåne County. In the simulation experiments, we assumed that all stroke patients have priority level one and need to be transported by ambulance to an acute hospital for diagnosis and treatment.

3.3 Configurations of simulation modeling framework and ambulance travel time estimation module

As mentioned earlier, the ANN model using the five selected features appeared to be the most effective for the travel time estimation problem. We trained the ANN model with the spatiotemporal data from Skåne County, incorporating these five features. We refer the reader to our companion paper [8] for details on the selection of ML models and hyperparameter configurations.

We illustrated the extended framework by constructing a simulation model for prehospital stroke care for Skåne County in Sweden. We conducted one simulation run using the constructed simulation model, both with and without making use of the ambulance travel time estimation module for travel time estimation. In the simulation model, we considered decision policies for 1) ambulance selection and 2) choice of destination hospital. For ambulance selection, we aimed to minimize the time to diagnosis by choosing the ambulance that is expected to reach the patient the fastest and then transfer them to the hospital as quickly as possible. For hospital selection, we opted to transport the patient to the nearest hospital. It should be noted that the time to diagnosis or treatment is the expected time from the onset of a stroke to when the patient receives diagnosis or treatment at the hospital.

In our scenario study, we assumed that all dispatches were conducted by ambulances, reflecting the current situation in the Skåne County. Additionally, we conducted another experiment to compare travel times for the simulated transport activities, with and without using the ambulance travel time estimation functionality.

3.4 Experimental results

In this subsection, we compare the performance of our EMS simulation modeling framework for stroke patients with and without the integration of the ML-based ambulance travel time estimation module. Without integration refers to using only dynamic travel time calculations, while with integration involves first calculating dynamic travel times, followed by estimation using the ML-based module. The emergency response metrics evaluated include time to diagnosis, time to treatment, and the time for three ambulance transport activities: ambulance-to-patient, patient-to-hospital, and hospital-to-hospital trips.

Table 1 presents the simulation results, comparing averages and medians for these emergency response metrics. The results show that integrating the ML-based estimation module leads to differences in emergency response metrics

Table 1. Comparison between the averages and medians for emergency response metrics, that is, time to diagnosis (TD), time to diagnosis for the secondary hospital (TD2), time to treatment (TT), time to treatment for the secondary hospital (TT2), and travel times from the ambulance location to the patient's location (A2P), from the patient's location to the hospital (P2H), and from one hospital to another (H2H), all presented in hours. The comparison is made between using the ML-based ambulance travel time estimation module (denoted as ML) and using only dynamic travel time without the ML module (denoted as Dynamic).

Metric	**Dynamic average**	**ML Average**	**Dynamic Median**	**ML Median**
TD	0.73	0.77	0.70	0.76
TT	1.32	1.35	1.28	1.35
A2P	0.18	0.22	0.16	0.21
P2H	0.25	0.25	0.22	0.24
H2H	0.55	0.45	0.39	0.36
TD2	1.80	1.71	1.83	1.85
TT2	2.39	2.30	2.42	2.43

compared to using only dynamic travel times. The differences between the metrics values with and without using the ML-based estimation module in simulation suggest that ambulance driving speeds differ from those of passenger cars. This highlights the importance of using specialized models like the ML-based estimation module, which accounts for the unique conditions under which ambulances operate. However, due to data limitations, the results do not demonstrate a clear improvement in the accuracy of estimated ambulance travel times. Instead, the observed differences indicate that the integration of ML-based estimations influences the outcomes in EMS simulations.

In Figure 4, we show how the estimated travel times for ambulance-to-patient and patient-to-hospital trips differ over distances with and without using the ML-based approach in our simulation model. In both scatter plots in Figure 4, we observe that there are variations in travel times relative to the distance when comparing the dynamic travel time calculation and the ML-based estimation approach. In particular, for longer trips (more than 20 km), the ML-based approach provides smaller travel times. For medium distances (between 5 to 20 km), the travel times for the dynamic travel time calculation and the ML-based approach tend to overlap, indicating that the ML model performs similarly to the dynamic approach within this range. However, the plots illustrate that the ML-based approach tends to estimate longer travel times when the distance is shorter (less than 5 km) compared to the dynamic travel time calculation.

It is interesting to observe how travel distances vary across different trips. The ambulance-to-patient trips range from 0.07 km to 39.15 km, and the patient-to-

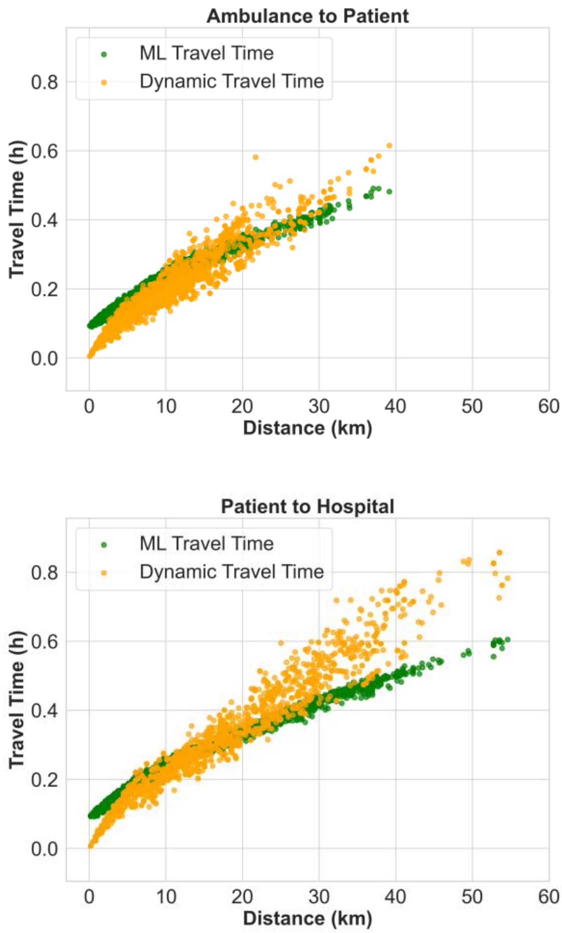


Figure 4. Scatter plots comparing dynamic travel time calculations and ML-based travel time estimations for ambulance-to-patient and patient-to-hospital trips.

hospital trips span from 0.18 km to 54.58 km. The hospital-to-hospital transport ranges from 0 km to 132.12 km, where 0 km indicates that no transport occurs.

We compared the execution time of the simulation model applied to Skåne County, indicating that incorporating the ML-based estimation module increased the execution time from 22.87 minutes to 28.66 minutes, a 5.79-minute increase. The increase is mainly due to applying the ML model to the travel times of approximately 3,000 synthetic stroke patients. For each patient, the dynamically calculated times for ambulance-to-patient, patient-to-hospital, and hospital-to-hospital trips are fed into the trained ANN model, adding extra processing time. While each call only takes a few milliseconds, the cumulative effect increases the total simulation time. While the additional computational burden caused by integrating the ML module may affect scalability, exploring how the ML-based approach influences the results can provide valuable insights.

The effectiveness of the ML-based estimation module is closely tied to the quality and quantity of data used for training the model. Although data from the Skåne County was limited, the results show that the ML model's ambulance travel time estimations differ in a reasonable way from those obtained using OSRM-CH, which is designed for passenger cars. These differences highlight the potential impact of ML-based estimations on simulation outcomes.

One limitation of our current spatiotemporal dataset is that it does not specifically focus on stroke patients or other high-priority emergency cases. Instead, it includes all types of ambulance travel activities, some of which may not be emergencies or high priority. This generalization might limit the ML-based model's effectiveness in estimating travel times for emergency cases. With more granular data, particularly data that differentiates between types of emergencies, priority levels, and contextual factors such as weather conditions, the ML-based model could be further refined to provide more accurate and robust estimations.

4. Conclusions

In the current paper, we elaborated on the idea of integrating the ML-based ambulance travel time estimation module proposed by Abid et al. [8] into an EMS simulation modeling framework. To assess the effectiveness of our approach, we built a simulation model for stroke patients and applied it in a scenario to the Skåne County, Sweden. The results demonstrated that the ML-based module affects the outcomes of emergency response metrics, showing differences in time to diagnosis, time to treatment, and ambulance transport activities compared to using only dynamic travel times. The observed differences highlight the potential influence of the ML-based estimations on the outcomes in EMS simulations. Additionally, incorporating the ML module did not add significant computational complexity, suggesting that it could be a viable addition to EMS simulation models.

References

[1] L. Aboueljinane, E. Sahin, and Z. Jemai, "A review on simulation models applied to emergency medical service operations," *Computers & Industrial Engineering*, vol. 66, no. 4, 2013.

[2] S. Amouzad Mahdiraji, M. Juninger, N. Narvell, J. Holmgren, R.-C. Mihailescu, and J. Petersson, "Implementing Dynamic Travel Time Calculation in EMS Simulations: Impacts on Prehospital Stroke Care and Transportation," in *Proc. Int. Conf. Health and Social Care Information Systems and Technologies (HCist)*, 2024.

[3] B. Hajinasab, P. Davidsson, J. Holmgren, and J. A. Persson, "On the use of on-line services in transport simulation," *Transportation Research Procedia*, vol. 21, pp. 208–215, 2017.

[4] B. Patel, N. M. Waters, I. E. Blanchard, C. J. Doig, and W. A. Ghali, "A validation of ground ambulance pre-hospital times modeled using geographic information systems," *Int. J. Health Geogr.*, vol. 11, no. 1, pp. 1–10, 2012.

[5] U. Mori, A. Mendiburu, M. Álvarez, and J. A. Lozano, "A review of travel time estimation and forecasting for advanced traveller information systems," *Transportmetrica A: Transport Science*, vol. 11, no. 2, pp. 119–157, 2015.

[6] N. Torres, L. Trujillo, Y. Maldonado, and C. Vera, "Correction of the travel time estimation for ambulances of the Red Cross Tijuana using machine learning," *Computers in Biology and Medicine*, vol. 137, pp. 104798, 2021.

[7] Ľ. Buzna and P. Czimmermann, "On the Modelling of Emergency Ambulance Trips: The Case of the Žilina Region in Slovakia," *Mathematics*, vol. 9, no. 17, pp. 2165, 2021.

[8] M. A. Abid, F. Lorig, J. Holmgren, and J. Petersson, "Ambulance travel time estimation using spatiotemporal data," *Procedia Computer Science*, vol. 238, pp. 265–272, 2024.

[9] T.-Y. Hu, C.-C. Tong, T.-Y. Liao, and W.-M. Ho, "Simulation-assignment-based travel time prediction model for traffic corridors," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 3, pp. 1277–1286, 2012.

[10] S. Amouzad Mahdiraji, J. Holmgren, A. Alshaban, R.-C. Mihailescu, J. Petersson, and J. Al Fatah, "A Framework for Constructing Discrete Event Simulation Models for Emergency Medical Service Policy Analysis," *Procedia Computer Science*, vol. 210, pp. 133–140, 2022.

[11] Statistics Sweden, "Demographic data 2023," [Online]. Available: https://www.scb.se/. [Accessed: Apr. 1, 2024].

[12] SOS Alarm, "SOS alarm data for 2020–2021," [Online]. Available: https://www.sosalarm.se/. [Accessed: Oct. 1, 2023].