8-2024

Application of Agent-Based Simulation and Game Theory in Evaluating Implementation of Whole Genome Sequencing in Treating Lung Cancer

Fateme Ghalenoei

THE UNIVERSITY OF NEW HAVEN

APPLICATION OF AGENT-BASED SIMULATION AND GAME THEORY IN EVALUATING IMPLEMENTATION OF WHOLE GENOME SEQUENCING IN TREATING LUNG CANCER

A THESIS

submitted in partial fulfillment

of the requirements for the degree of

MASTER OF SCIENCE IN INDUSTRIAL ENGINEERING

Advisor: Dr. Marzieh Soltanolkottabi

BY

Fateme Ghalenoei

University of New Haven

West Haven, Connecticut

August 2024

APPLICATION OF AGENT-BASED SIMULATION AND GAME THEORY IN EVALUATING IMPLEMENTATION OF WHOLE GENOME SEQUENCING IN TREATING LUNG CANCER

SUBMITTED BY

Fateme Ghalenoei (Aug 23, 2024 10:15 CDT)

[Fateme, Ghalenoei]

APPROVED BY

Marzieh Soltanolkottabi (Aug 23, 2024 11:43 EDT)

[Marzieh Soltanolkottabi, Ph.D.]

Thesis Adviser

Ali Montazer

[M.Ali Montazer, Ph.D.]

Committee Member

[Narjes Sadeghiamirshahidi, Ph.D.]

Committee Member

Adwoa Donyina (Aug 29, 2024 17:47 EDT)

[Adwoa Donyina, Ph.D.]

Committee Member

Nadiye Erdil

[Nadiye O. Erdil, Ph.D.]

Program Coordinator

RSHarichandran

[Ronald S. Harichandran, Ph.D.]

Dean of the College of Engineering

Nancy D Savage

[Nancy Savage Ph.D.]

Provost

Acknowledgements

Sincerest thanks to Dr. Marzieh Soltanolkottabi, Professor M Ali, Montazer, Dr. Narjes, Sadeghiamirshahidi and Dr. Donyina for their patience and support.

Special thanks also to my parents; and appreciation to UNH.

Dedication

To my beloved parents, for their endless love, support, and sacrifices that made this journey possible. And to my cherished cousin who suffered from muscular dystrophy. Mohtada, who "flew like the birds and ascended to the highest heavens." Though you are no longer with us, your memory is always with us.

ABSTRACT

Cancer, particularly lung cancer, presents significant diagnostic and economic challenges globally. Timely diagnosis and cost management play pivotal roles in treatment success. A biomarker is any measurable molecule in blood, bodily fluids, or tissues, indicating the potential presence of an abnormal bodily process, condition, or disease. Biomarker testing is a laboratory test in oncology that is used in the selection of targeted cancer treatments and to help avoid ineffective treatments. Whole Genome Sequencing (WGS), is a biomarker test which while more comprehensive, comes at a higher cost. This study proposes an agent-based simulation model within a game-theoretic framework to examine the benefits of prioritizing WGS in the diagnostic process for lung cancer across various hospital settings. The typical diagnostic pathway for lung cancer, involving tests like PD-L1, ALK, EGFR, KRAS, BRAF, and ROS1, can lead to time-consuming referrals and potentially higher costs due to unsuccessful standard of care (SOC) tests. This thesis scrutinizes these pathways and evaluates the impacts of different referral scenarios on the number of diagnosed patients and treatment costs, drawing on a system dynamics model. Initial findings suggest that uniform referral strategies are not universally optimal and can result in delayed WGS testing results. Aiming to establish a refined strategy for each hospital type and determine the optimal timing for WGS, an agent-based simulation is proposed to emulate the diagnostic journey. The model considers costs, success rates, and diagnostic timeframes, while a game theoretic approach assesses each hospital's decision-making regarding WGS access. The goal is to facilitate the earliest and most cost-effective treatment for patients. Objectives of this thesis include developing an agent-based model to mirror current pathways, evaluating referral scenarios, formulating a game-theoretic analysis, and understanding the sensitivity of various parameters crucial to decision-making. The research will culminate in actionable insights for healthcare

providers and policymakers, specifically in enhancing WGS utilization in lung cancer diagnosis and treatment.

Anticipated outcomes include a versatile model for strategizing referral practices, yielding an optimal approach for a given health center constellation. The broader implications of this work extend beyond lung cancer, offering a template for enhancing diagnostic efficiency in various cancers and other sectors where strategic decisions regarding shared resources are vital. The study stands to influence patient care positively by improving diagnostic accuracy and economic efficiency in healthcare delivery.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: INTRODUCTION

1.1. Statement of the Problem

Cancer is one of the most fatal diseases which annually affects many individuals all around the world. The challenges of cancer treatment are not limited to the treatment process itself. The timely diagnoses of the patients and the costs involved have a crucial role in the success of the treatment efforts. Biomarker testing is a laboratory test in oncology that is used in the selection of targeted cancer treatments to help to avoid ineffective treatments. Whole Genome Sequencing (WGS) is a biomarker test for analyzing the entire genome which can provide more comprehensive diagnostic information. WGS is more expensive than other tests and not all the hospitals provide it. For lung cancer patients, many hospitals offer a standard of care (SOC) that encompasses certain biomarker tests such as PD-L1, ALK, and EGFR. If a patient cannot be diagnosed using SOC, they will be referred to other hospitals to be tested for other biomarker tests or will be referred for WGS testing. These referrals are usually based on the patient's proximity to hospitals. Patients are referred to Teaching hospitals if biomarkers cannot be identified in General hospitals, and to Academic hospitals for WGS testing if biomarkers cannot be identified in Teaching hospitals. Unfortunately, patients often waste their time waiting for the results of SOC testing when they could be referred for WGS testing. In some cases, the cost of performing unsuccessful biomarker tests is even greater than the cost of WGS testing for patients [1]. Depending on their proximity to WGS centers, different hospitals may require distinct referral strategies. Moreover, if all hospitals choose to refer all their patients for WGS testing, it could potentially cause delays in receiving test results. Therefore, in order to establish an optimal strategy for each hospital and determine the most appropriate timing for conducting WGS testing, we propose utilizing an agent-based simulation to

simulate the diagnostic process. To enhance the hospital's choice, we will adopt a game theoretic framework. By allowing each hospital to develop its own diagnostic policy and approach for using WGS, patients can receive treatment at the earliest possible time and at the lowest possible cost.

1.2 Methods of Research to be Utilized

Agent-based simulation (ABS) and game theory are computational modeling techniques that can be used to study complex systems and interactions between different agents. In the case of cancer treatment, these techniques can be used to model the timing of implementation of WGS testing to identify the optimal strategy for different hospitals. By creating a simulation that includes different types of hospitals at different locations with varying levels of testing capabilities, we can model the interactions between hospitals and the decision-making processes involved in referring patients for WGS testing. In the simulation model we will consider different parameters including the cost of testing, success rates of tests, and the time taken to complete the diagnostic pathway. The data for the parameters will be derived from available online resources [2]. We will analyze the simulation output to derive meaningful insights for the following purposes:

- Evaluate various referral scenarios in terms of associated costs and time.
- Investigate the individualized timing for the implementation of WGS testing in different hospitals.
- Examine the model's sensitivity to changes in various parameters and identify the most critical parameters for decision-making.

1.3 Way in Which the Research Will Contribute an Original Point of View

This research represents a groundbreaking contribution to the field of healthcare management and genomics implementation by introducing a novel perspective on enhancing diagnostic processes

for lung cancer patients. The innovative aspect lies in its exploration of the possibility that different hospitals may require distinct referral strategies based on their proximity to WGS centers. Additionally, the proposal to employ an agent-based simulation within a game theoretic framework to model and evaluate hospital-specific diagnostic policies is a pioneering approach. This study has the potential to improve the diagnosis and treatment of lung cancer patients by providing recommendations on the optimal use of WGS testing. The proposed model can also be adapted and applied to other types of cancer, providing a framework for optimizing the diagnostic pathway for different types of cancer. Furthermore, the findings of this study can be implemented in any type of problem in which multiple agents must make decisions about utilizing a more efficient shared service where time and cost are key factors in the decision-making process. In a business context, multiple small businesses may choose to outsource certain operations, but as more businesses opt to outsource, the waiting time for services may increase.

CHAPTER II: LITERATURE REVIEW

2.1 Whole Genome Sequencing

2.1.1 An Introduction to Whole Genome Sequencing

WGS is a laboratory technique designed to examine complete genomes in their entirety [3]. In molecular biology and genetics, a genome encompasses the complete genetic data of an organism, comprising the extensive collection of genetic directives. Within every genome lies the entirety of instructions required for constructing the organism and facilitating its growth and maturation [4]. WGS is an advanced and thorough genetic analysis method that acquires comprehensive data on every gene and all the chromosomes in DNA [5]. WGS reads and decodes the complete genetic instructions (DNA) in an organism's genome. This process provides a detailed map of an individual's genetic information, including all the genes and non-coding regions, helping researchers and healthcare professionals understand genetic variations, potential health risks, personalized healthcare options and more. Here's a more detailed explanation of whole genome sequencing:

DNA Sequencing: WGS is a laboratory process that reads the order of the four DNA bases (adenine, thymine, cytosine, and guanine) in an individual's genome. This process provides a precise blueprint of the individual's genetic code.

Comprehensive Analysis: Unlike targeted genetic tests that focus on specific genes or regions, WGS aims to capture every part of the genome, from coding regions that produce proteins to non-coding regions with regulatory functions.

Variation Detection: WGS is valuable for identifying genetic variations, including single nucleotide changes (SNPs), insertions, deletions, and structural variations. These variations can be linked to genetic diseases, susceptibility to certain conditions, and even ancestry information.

Clinical Applications: WGS has important clinical applications, including diagnosing genetic disorders, identifying potential treatment options, and understanding the genetic basis of diseases. It can help tailor personalized medicine approaches based on an individual's unique genetic makeup.

2.1.2 Understanding WGS as a Cancer Biomarker Test

WGS serves as a biomarker test dedicated to the analysis of the entire genome [1]. To explain the biomarker test, first we need to answer the question "What is biomarker?" As defined by the National Cancer Institute (NCI), a biomarker is a biological molecule present in blood, other bodily fluids, or tissues. It serves as an indicator of a regular or anomalous biological process, or as a signal of a specific disorder or disease, such as cancer [6].

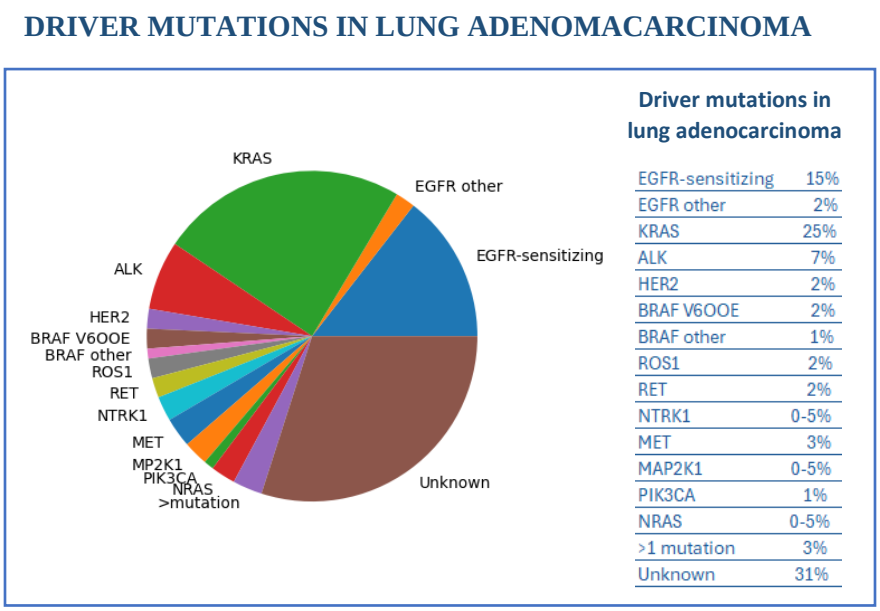

Figure 1: Prevalence of Driver Mutations in Lung Adenocarcinoma

Figure 1 is inspired by an image from the website, but modifications have been made to fit the context of this thesis. [7]

A biomarker test, also referred to as a biomarker assay, is a medical examination that assesses specific molecules, genes, proteins, or other biological markers within an individual's body. These

biomarkers offer crucial insights into one's health, the presence of a disease, its evolution, and how it responds to treatment. Biomarker tests play a pivotal role across a spectrum of medical disciplines, encompassing cancer diagnosis and staging, treatment selection, monitoring disease progression, cardiovascular health, infectious disease detection, tracking drug levels, genetic analysis, and the study of neurological disorders. In this research, my primary emphasis has been on the field of cancer diagnosis., where cancer biomarkers—whether biological, chemical, or biophysical—reside in tumor tissues or bodily fluids, providing essential information about the current and future behavior of cancer [8].

2.1.3 Genetic Mutations in Cancer

Cancer, fundamentally a genetic disorder, emerges from genetic mutations and alterations in transcriptional and epigenetic patterns. These genetic transformations offer vital insights as biomarkers for the early identification, staging, and in-depth molecular analysis of cancer, enabling tailored therapies. Well-recognized oncogenes such as EGFR, HER2, and KRAS, alongside tumor suppressor genes like TP53, PTEN, and PI3K, are currently employed as biomarkers to direct treatment strategies in various cancer types, including breast cancer, ovarian cancer, lung cancer, and prostate cancer, among others [9].

There are different types of genetic mutations based on where they form.

Types of genetic mutations include:

Germline mutation: A change in a gene taking place within a parent's reproductive cells (egg or sperm) can influence the hereditary traits of their offspring.

Somatic mutation: A gene alteration that takes place post-conception within the developing embryo, potentially leading to a future baby, is termed a somatic mutation. Such mutations

manifest in all cells of the developing body, except for the sperm and egg. It's important to note that somatic mutations cannot be inherited from parents to their children, as hereditary traits are exclusively transmitted through the sperm and egg.

Numerous genetic conditions are known to exist, with some of the most prevalent including Alzheimer's disease, some cancers, cystic fibrosis, Down syndrome, and sickle cell disease. It's noteworthy that mutations play a significant role not only as a distinguishing feature of cancer but also in shaping the evolution of cancer itself [10].

2.1.4 Why Do We Use Biomarker Test?

Cancer encompasses a diverse group of diseases marked by the accelerated proliferation of abnormal cells that extend beyond their typical confines. This, in turn, can lead to the dissemination of tumor cells to distant organs and regions of the body, a process known as metastasis. In 2018, cancer claimed approximately 9.6 million lives, making it the second most prevalent cause of death globally. Consequently, it remains a crucial and indispensable domain for research and investigation [10]. The early detection of cancer and the efficient management of treatment costs represent two key hurdles. As previously discussed, biomarker tests play a crucial role in cancer diagnosis. Among these, WGS stands out as a biomarker test offering more extensive diagnostic insights, albeit at a higher cost compared to other tests. Moreover, it allows for the simultaneous examination of multiple genes within a single test, thereby saving valuable time—a critical factor in the realm of cancer treatment [11]. WGS offers greater advantages when compared to other biomarker tests, albeit at a higher cost than other tests [1].

2.1.5 Non-Small Cell Lung Cancer

This study centers on lung cancer, a malignancy that originates in the lungs and has the potential to metastasize to lymph nodes and other body organs, including the brain. Lung cancers are typically categorized into two primary types, namely small cell and non-small cells, with non-small cells encompassing subtypes such as adenocarcinoma and squamous cell carcinoma [12]. Patients diagnosed with stage IV non-small-cell lung cancer (NSCLC) must undergo a series of biomarker tests to facilitate an accurate diagnosis and receive the most suitable treatment. Among these tests, WGS emerges as a potent tool for biomarker identification. Early integration of WGS into the cancer treatment process streamlines the diagnostic phase, reducing the need for multiple tests and expediting the initiation of treatment. However, it's essential to acknowledge that WGS comes at a higher cost when compared to individual biomarker testing and is not available in all types of hospitals. Consequently, some hospitals may need to send samples to specialized centers, which can introduce additional delays. Notably, while WGS has the capacity to identify a broader spectrum of biomarkers, not all patients necessarily require this comprehensive testing. In many instances, standard diagnostic procedures suffice for biomarker identification. The challenge lies in determining the optimal timing for WGS testing to enable patients to start treatment at the earliest possible juncture and with the most cost-effective approach [1].

2.1.6 Cost of Whole Genome Sequencing

In Table 1, the expenses associated with genome sequencing are presented. The overall cost for sequencing the genome in a cancer case amounts to £6,841, with each genome costing £3,420 [13]. The study took place in Brisbane, Australia, spanning 2017 and 2018, and it employed a combined

approach of gross and micro-costing methods to compile comprehensive resource data for a total of 1,433 patients. The research investigated the expenses associated with genomic sequencing for patients diagnosed with lung, breast, esophageal cancers, melanoma, or mesothelioma. These costs were categorized into seven distinct activities within the sequencing process, namely sampling, extraction, library preparation, sequencing, analysis, data storage, and clinical reporting. The findings of the study revealed that the per-person cost for whole genome sequencing ranged from AU$2,895 to AU$4,830, equivalent to US$2,006 to US$3,347 in 2018 [14].

Table 1 : Overall cost for sequencing the genome.

Stage	Cost Category			Total	Total test costs before overheads
	Equipment	Consumables	Staff		
Sample reception	€ 0.24	€ 0.31	€ 17.59	€ **18.14**	0.3%
DNA extraction	€ 0.19	€ 20.91	€ 15.95	**37.06**	0.7%
Nanodrop	€ 0.16	€ 0.11	€ 6.20	€ **6.47**	0.1%
Qubit	€ 0.19	€ 3.05	€ 11.20	€ **14.43**	0.3%
Agarose gel	€ 0.15	€ 14.26	€ 22.70	€ **37.11**	0.7%
*Library processing	€ 51.83	€ 132.09	€ 84.64	€ **268.56**	4.7%
**Sequencing	€ 615.24	€ 3,688.04	€ 48.55	€ **4,351.83**	76.3%
Bioinformatics	€ 1.80	€ 266.73	€ 406.99	€ **675.52**	11.8%
Reporting	€ -	€ -	€ 257.25	€ **257.25**	4.5%
Data archiving	€ 24.56	€ 0.96	€ 8.83	€ **34.35**	0.6%
Total (before overheads)	€ 694.35	€ 4,126.46	€ 879.89	€ **5,700.71**	_
% total cost	12%	72%	15%	_	_
Total (including overheads calculated at 20%)	_	_	_	**6840.85 (3420.43 per genome)**	_

*Library processing includes library preparation, normalization and validation.

*Sequencing also includes clustering.

2.1.7 Whole Genome Sequencing in the United States

WGS is not widely used in the USA and is currently offered to children hospitalized with severe illnesses. This service is made possible through a collaboration between clinical professionals and laboratory genomics experts from Pediatrics, Medical Genetics, and the Yale Center for Genome Analysis. Presently, WGS is available to patients in the neonatal intensive care unit (NICU) and the pediatric intensive care unit (PICU) at Yale New Haven Children's Hospital [15]. Since clear data for the US is not clearly available and the US health system is complex, the data from the Netherlands is used instead.

2.2 Agent-Based Simulation and Game Theory

2.2.1 What is Agent-Based Simulation?

Agent-based simulation (ABS) is an innovative approach for modeling complex systems, comprising autonomous, interacting "agents". This methodology facilitates the modeling of individual behaviors and their impact on others in ways not previously achievable. ABS is a relatively recent method for simulating intricate systems composed of interacting, self-governing "agents." These agents exhibit behaviors typically governed by simple rules and engage in interactions with one another, leading to the modification of their behaviors. By modeling agents individually, we gain insights into the diversity present in their attributes and behaviors, which collectively influence the overall system behavior. A bottom-up approach to modeling, agent by agent, and interaction by interaction, often reveals instances of self-organization, where unprogrammed patterns, structures, and behaviors emerge as a result of agent interactions. Emphasizing the modeling of agent heterogeneity and the emergence of self-organization distinguishes agent-based simulation from other simulation techniques such as discrete-event simulation and system dynamics. Agent-based modeling provides a powerful means to simulate social systems composed of agents that interact, influence each other, learn from experiences, and adapt their behaviors to better suit their environment [16].

2.2.2 Structure of an Agent-Based Model

A typical agent-based model consists of three key components:

1. Agents, their characteristics, and their actions.
2. Agent relationships and ways they interact: The structure of connections specifies how agents interact with each other.
3. The environment where agents operate, in addition to interacting with other agents.[17]

2.2.3 Agent-Based Modeling Applications

Agent-based modeling has found diverse applications in fields such as physical, biological, social, and management sciences. These applications encompass a wide spectrum, from simulating agent behavior in supply chains and financial markets to predicting the outcomes of marketing campaigns and the spread of diseases, as well as forecasting the future healthcare system needs. The utility of agent-based models extends from re-creating ancient civilizations that vanished centuries ago to devising models for entirely new markets that are yet to exist. These models come in various forms, ranging from simple academic models that capture the core aspects of a system to sophisticated decision support systems used in real-world scenarios. Minimalist models make use of simplified assumptions to focus on the essential system features, while decision support models are tailored for practical policy questions, incorporate real data, and undergo rigorous validation to ensure their reliability [16].

2.2.4 Agent-Based Simulation in Healthcare

In the last 15 years, ABS as experienced significant growth across various fields, including hospital and healthcare settings. In healthcare, ABS finds valuable applications in studying patient flow, such as in emergency departments, and addressing operational challenges within hospitals. Additionally, ABS has been utilized to investigate the spread of infections within hospital environments, like understanding the hospital's role during influenza epidemics and analyzing the dynamics of nosocomial infections. Economic healthcare models that don't focus on individual patient scale have also been explored using ABS, although these models are not covered in this survey [18].

ABS is quickly gaining popularity in various fields. It complements other simulation methods like system dynamics and discrete event simulation. ABS works especially well in healthcare

operations management due to the complex interactions among individuals. ABS explicitly models these interactions, uncovering system behaviors not visible through other methods. ABS has found success in various aspects of healthcare operations management, including healthcare delivery, epidemiology, economics, and policy [19].

2.2.5 An Introduction to Game Theory

Game theory delves into mathematical models that explore conflict and cooperation among intelligent and rational decision-makers. To grasp its core concepts, it's helpful to start by delving into decision theory. Decision theory centers around the predicament faced by an individual tasked with selecting from a range of uncertain options often referred to as "lotteries." These lotteries yield outcomes or "prizes" determined by chance, sometimes influenced by unknown factors, which we can label as the "state" or "state of the world." Game theory is all about examining mathematical models that deal with conflicts and collaborations involving smart, logical decision-makers. When we say "rational," we mean that each person's choices aim to maximize their expected outcomes, assuming they know what the others are up to. And by "intelligent," we mean that everyone involved comprehends the setup and that everyone else is also a smart, rational decision-maker. So, when we come up with a theory describing how players should act in certain games, we assume that each player in the game also knows and follows this theory and its predictions [20]. Game theory is the science of managing conflicting interests. Conflict arises when two or more opposing "individuals" must make decisions that lead to various possible outcomes, depending on what their opponents decide. These individuals hold preferences that clash with those of their opponents. Game theory dissects these conflicts, outlines the choices made by everyone, and scrutinizes the potential results of the competitive game.

Game theory is a mathematical framework used for analyzing situations where multiple decision-makers, called players, interact strategically. It aims to understand the behavior of these players and predict the outcomes of their interactions based on their preferences, strategies, and possible actions [21].

2.2.6 Spatial and Evolutionary Game Theory

Spatial game theory is a branch of game theory that focuses on the spatial aspects of strategic interactions among players. Unlike traditional game theory, which typically assumes players interact uniformly, spatial game theory considers the locations and spatial distribution of players, often modeling interactions on grids or networks where each player interacts primarily with their neighbors. This approach is particularly useful for studying phenomena where the spatial arrangement affects outcomes, such as the spread of diseases, the evolution of cooperation, or the competition among firms in a market. By incorporating spatial dimensions, this theory provides insights into how local interactions can lead to complex global patterns and behaviors, revealing the importance of spatial structure in strategic decision-making processes [22].

Evolutionary game theory provides a framework to study the dynamics of strategic interactions in populations over time. Unlike traditional game theory, which often assumes rational players with complete knowledge and perfect foresight, evolutionary game theory focuses on the adaptation and evolution of strategies within a population of boundedly rational agents. One of the classic models used in evolutionary game theory is the repeated Prisoner's Dilemma game, which provides insights into the evolution of cooperation and defection strategies. The field of evolutionary game theory, such as the work by Weibull in 1995, has emerged by integrating game theory with the fundamental principles of Darwinism. This fusion aims to address the temporal aspect that the original game theory lacks, as it predominantly deals with equilibrium [23]. Spatial evolutionary

games represent a timeless and extensively examined model for exploring evolutionary dynamics on graphs. This model has garnered attention and investigation across a wide spectrum of disciplines, including biology, physics, and computer science [24].

2.2.6.1 The Repeated Prisoner's Dilemma Game

The Prisoner's Dilemma is a fundamental model in game theory that demonstrates the conflict between individual rationality and collective welfare. In a single iteration of the game, two players can choose to either cooperate or defect. The payoffs are structured such that mutual cooperation yields a moderate reward for both players, mutual defection results in a low payoff, and if one player defects while the other cooperates, the defector receives the highest payoff while the cooperator receives the lowest. In the repeated Prisoner's Dilemma, the game is played multiple times, allowing players to potentially adjust their strategies based on the outcomes of previous rounds. This repetition enables the exploration of strategy evolution, as players may develop strategies based on past interactions, such as tit-for-tat, always cooperate, always defect, or more complex strategies.

To illustrate the concept, consider the following payoff matrix for a single iteration of the Prisoner's Dilemma:

Table 2: Payoff Matrix

	Player B Cooperates	**Player B Defects**
Player A Cooperates	(3, 3)	(0, 5)
Player A Defects	(5, 0)	(1, 1)

In this matrix:

- If both players cooperate, they each receive a payoff of 3.

- If player A cooperates while Player B defects, Player A receives a payoff of 0 while Player B receives 5.
- If Player A defects while Player B cooperates, Player A receives 5 and Player B gets 0.
- If both players defect, they each receive a payoff of 1.

This setup highlights the dilemma: mutual cooperation leads to a better collective outcome than mutual defection, but individual rationality drives players towards defection, as it yields a higher individual payoff when the other player cooperates.

2.2.6.2 The Repeated Prisoner's Dilemma

In the repeated Prisoner's Dilemma, the game is played multiple times, allowing players to potentially adjust their strategies based on the outcomes of previous rounds. This repetition enables the exploration of strategy evolution, as players may develop strategies based on past interactions, such as tit-for-tat, always cooperate, always defect, or more complex strategies.

2.2.6.3 Payoff Calculation

In evolutionary game theory, the fitness of a strategy is determined by its payoff, which is typically the average reward received from interactions with other members of the population. In the repeated Prisoner's Dilemma, the payoff of a player is the sum of the payoffs received in each round, averaged over the number of rounds played. The payoffs depend not only on the player's own strategy but also on the strategies of others in the population. For instance, consider a population consisting of cooperators and defectors. If the majority of the population consists of defectors, the average payoff for defectors will be higher since they exploit cooperators. However, if the population has a significant number of cooperators, a cooperator might achieve a higher payoff by cooperating with another cooperator, making cooperation a viable strategy.

2.2.6.4 Strategy Update Mechanisms

In evolutionary repeated games, strategy updates occur as players adjust their strategies based on the payoffs they receive, often modeled through imitation or learning mechanisms. One common mechanism is the **imitate the best** rule, where players adopt the strategy of the most successful individuals in the population. This rule assumes that players have limited information and simply copy the strategy of those they observe to be doing well.

Another approach is the **proportional imitation rule**, where the probability of adopting a particular strategy is proportional to its relative success. In this mechanism, players are more likely to adopt strategies that yield higher payoffs, but there is also some probability of retaining or experimenting with less successful strategies, allowing for diversity in the population.

Additionally, the **replicator dynamics** is a mathematical model used to describe how the proportion of strategies evolves over time. According to this model, the growth rate of a strategy's frequency in the population is proportional to the difference between the average payoff of the strategy and the population's average payoff. Strategies that perform better than average grow in frequency, while those that perform worse decline [25].

2.2.7 Application of Agent-Based Simulation in this study

In this study, ABS and game theory are applied to evaluate the referral strategies of different hospitals for Whole Genome Sequencing (WGS) testing in lung cancer diagnosis. The agent-based model will simulate interactions between hospitals, incorporating various parameters such as testing costs, success rates, and diagnostic timeframes to mimic real-world decision-making processes. Game theory will then be used to analyze these strategies, treating each hospital as a strategic player to determine the most cost-effective scenario for referring patients for WGS

testing. This integrated approach aims to develop effective diagnostic policies that minimize costs and maximize the number of diagnosed patients.

2.3 Conclusion

Chapter II provides a comprehensive review of WGS, its applications, and its significance in the field of cancer diagnosis. WGS offers a detailed map of an individual's genetic information, enabling the detection of genetic variations that can be linked to various diseases, including cancer. The use of WGS as a biomarker test provides crucial insights into the genetic makeup of tumors, facilitating personalized treatment approaches. Furthermore, the chapter discusses the practical aspects of implementing WGS in clinical settings, including its cost and availability. While WGS is not widely used in the USA due to the US complexity in insurance, its potential for comprehensive diagnostic insights makes it a valuable tool in advanced healthcare settings. Additionally, the chapter introduces ABS and game theory as innovative approaches for modeling complex systems in healthcare. These methods enable the simulation of interactions between hospitals and the assessment of referral strategies for WGS testing in lung cancer diagnosis. By integrating ABS and game theory, the study aims to develop effective diagnostic policies that minimize costs and increase the number of diagnosed patients. This literature review sets the foundation for the subsequent chapters, which will explore the practical application of these methodologies to enhance lung cancer diagnosis and treatment.

CHAPTER III: SIMULATING WHOLE GENOME SEQUENCING IN LUNG CANCER DIAGNOSIS PROCESS

3.1 An Agent-Based Simulation

ABS is a technique that can be used to study complex systems and interactions between different agents. When examining cancer treatment, this tool can help determine the optimal timing for implementing WGS testing in various hospitals. By creating a simulation that includes different types of hospitals at different locations with varying levels of testing capabilities, we can model the interactions between hospitals and the decision-making processes involved in referring patients for WGS testing. Parameters such as test success rates, cost and the time required for the diagnostic pathway are carefully considered in the simulation model.

3.1.1 Building Models with Mesa

In this study, the Mesa framework was utilized to build agent-based models due to its robust and versatile platform for developing complex simulations. Mesa is an open-source Python library specifically designed for agent-based modeling. The choice of Mesa was driven by its ability to handle the intricate dynamics of agent interactions effectively, its strong community support, and its seamless integration with other scientific libraries in Python, such as NumPy and Matplotlib. These capabilities make Mesa an ideal tool for modeling and analyzing the intricate processes and interactions within the study, ensuring both accuracy and efficiency in simulation development.

3.2 Problem description

3.2.1 Diagnostic Pathways and Whole Genome Sequencing in Stage IV NSCLC Treatment

The cancer type addressed in this thesis is lung cancer, and the setting under consideration is the current practice in the Netherlands [2]. Generally, upon diagnosis of stage IV NSCLC (Non-Small-Cell Lung Cancer), patients undergo a series of biomarker tests to complete their diagnostic

pathway before initiating treatment. The assumption is that patients initially seek care at the nearest hospital, leading to different pathways based on the type of hospital visited: General, Teaching, or Academic. Among these, only Academic hospitals offer WGS facilities. General hospitals primarily conduct simpler tests targeting PD-L1, ALK, EGFR, and KRAS genes, while Teaching hospitals expand the testing scope to include ALK, PD-L1, KRAS, BRAF, EGFR, and ROS1. Academic hospitals utilize the same standard-of-care (SOC) tests as teaching hospitals but extend their services to include WGS testing for referred patients. Patients interact with different types of hospitals based on their proximity. When patients visit General or Teaching Hospitals, biomarker testing is conducted to diagnose their condition. If the testing is successful, indicating a positive diagnosis, the patient exits the system. However, if the testing is unsuccessful, suggesting a need for further evaluation, the patient is referred to a more specialized hospital. At Academic Hospitals, patients have additional options for diagnosis. Here, there's a chance to utilize SOC procedures or undergo WGS. The decision to refer a patient for WGS biopsy depends on various factors, including the willingness of both the patient and the physician. Once a WGS biopsy is performed, the number of tumor cells in the sample is examined. If the percentage of tumor cells is low, indicating a less severe condition, the patient would typically be referred back to SOC. However, if the biopsy is suitable for WGS analysis, it is sent to the WGS facility for further examination and the preparation of a detailed report. This report is then forwarded to a Multidisciplinary Tumor Board (MTB) for interpretation. Each pathway within this diagnostic process incurs costs and requires a certain amount of time. These outcomes, including cost and time to diagnosis, are essential metrics used to evaluate different scenarios and optimize the healthcare system's efficiency [1].

3.2.2 Patient Referral Process Based on Proximity

Patients who visit general, teaching, or academic hospitals on their own go to the hospitals closest to them. For example, patient 1 chooses to go to general hospital 1, which is the closest among the six general hospitals. If patient 1 is not diagnosed at general hospital 1, they are supposed to be referred to one of the teaching hospitals. General hospital 1 then sends the patient to the teaching hospital that is closest to it.

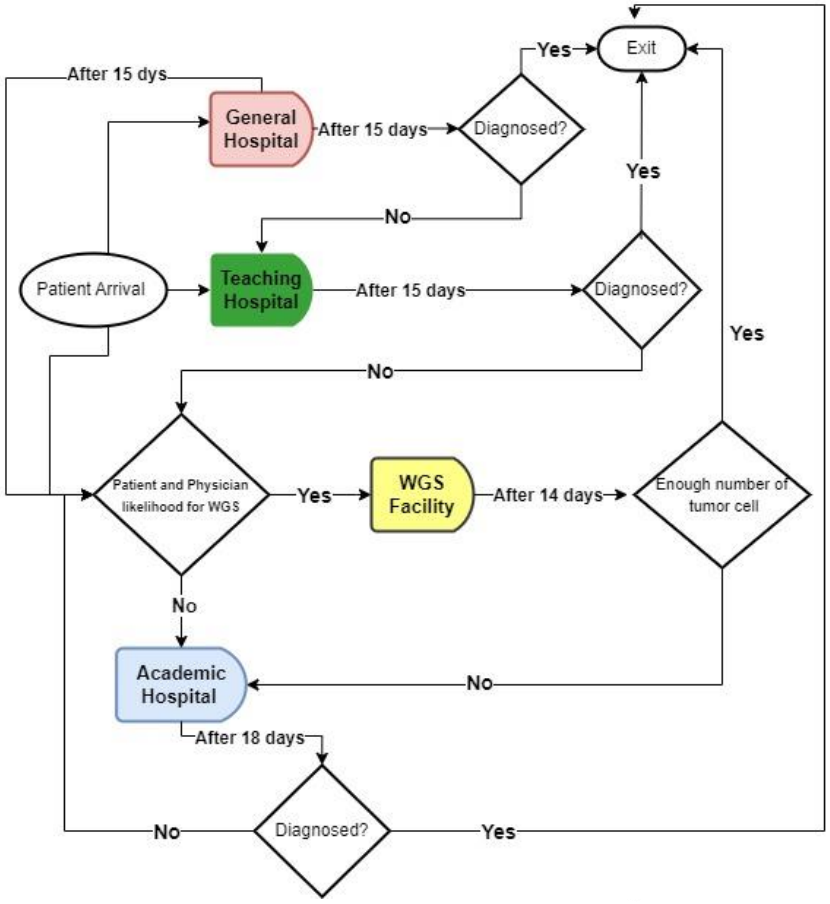

Figure 2: Flowchart Depicting the Diagnostic Referral Pathways for NSCLC Patients at Various Hospital Levels

This flowchart illustrates the patient's diagnostic process. Patients visit hospitals. If a patient has been diagnosed at a general hospital, they will exit the system and be considered as diagnosed patients. If not diagnosed, they will be referred to a teaching hospital for further tests. If diagnosed at a teaching hospital, they will exit the system as diagnosed patients; otherwise, they will be

referred to an academic hospital. At the academic hospital, based on the patient and physician's willingness, WGS will be conducted. Once a WGS biopsy is performed, the number of tumor cells in the sample is examined. If the percentage of tumor cells is low, indicating a less severe condition, the patient would typically be referred back to SOC.

3.3 Diagnostic Pathways

In the typical diagnostic process, if a patient initiates their diagnostic pathway at a general hospital and their biomarkers remain unidentified, they are referred to a teaching hospital for further testing. If the patients cannot be diagnosed at teaching hospital and identification of the biomarkers fails, they will be referred to academic hospitals for WGS testing. Patients referred for WGS testing are consulted for approval before the process begins. Academic hospitals send biopsy samples to WGS facilities for testing, and upon completion, the results are forwarded to the MTB for interpretation. If WGS testing fails, patients revert to receiving SOC treatment. Throughout this process, referrals are determined based on geographical proximity.

3.3.1 Suggested Referral Pathway: Direct Transfers from General to Academic Hospitals

We suggested a scenario where patients are referred directly to an academic hospital from a general hospital. In the typical diagnostic process, patients visit a general hospital and are then referred to a teaching hospital. From there, they may be sent to an academic hospital. In the suggested scenario, the process remains largely the same, except that patients are directly referred to an academic hospital from the general hospital, bypassing the teaching hospital entirely. "Bypassing the teaching hospital entirely" means that in the suggested scenario, patients are sent directly from a general hospital to an academic hospital, without being referred to or treated at a teaching hospital in between these steps. If we consider the typical diagnostic process as the base scenario and refer

to the alternative as Scenario 2, both base scenario and scenario 2 have been modeled in Python using the Mesa package to compare the differences between them.

3.3.2 Models' Assumptions

Three different types of hospitals have been considered: General Hospital, Teaching Hospital, Academic Hospital, and WGS Facility. Each of these is regarded as an agent, with its own attributes and behaviors.

Here are the agents with their attributes:

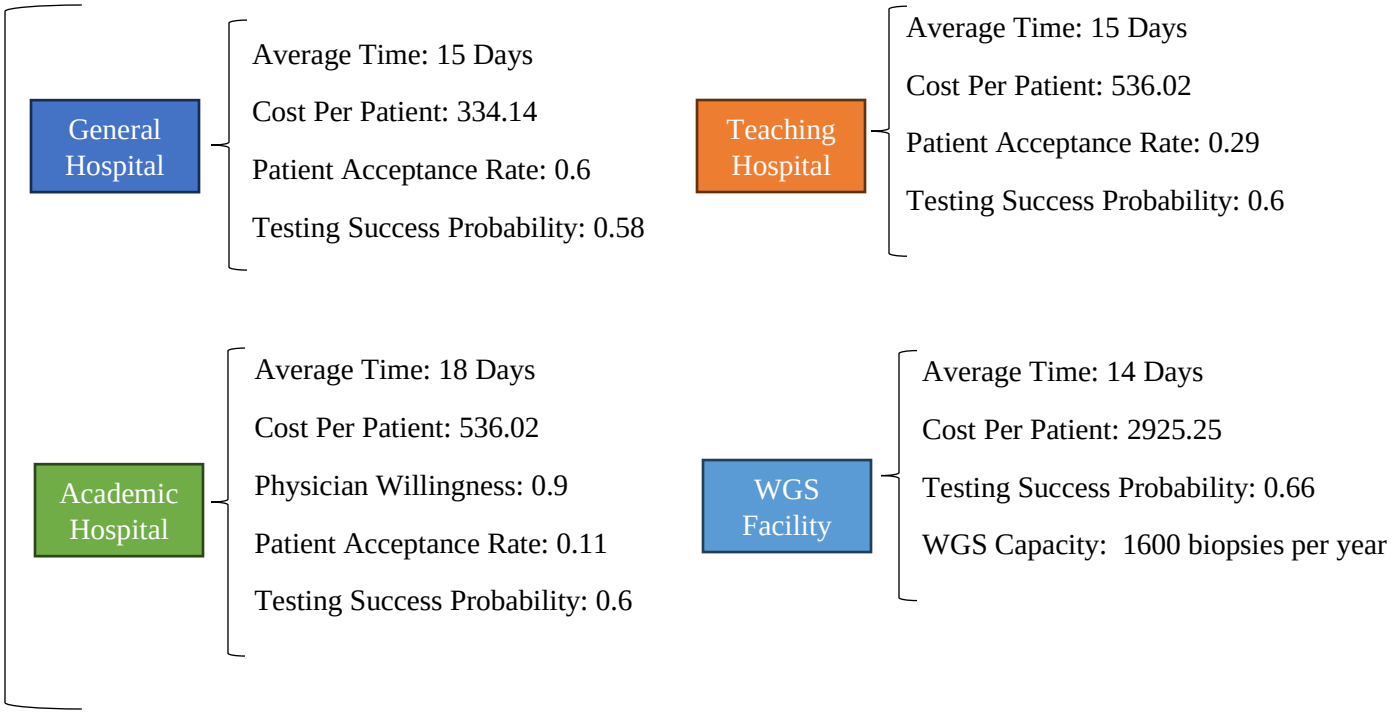

Figure 3: Agent's attributes.

These numbers which are written in figure 3 are derived from a study published in [1].

Mesa has two classes: Agent class and model class. The model simulates a healthcare system using agent-based modeling techniques. The healthcare system comprises multiple types of hospitals, including General Hospitals, Teaching Hospitals, and Academic Hospitals, as well as a Whole

Genome Sequencing (WGS) Facility. The model aims to explore patient flow, diagnosis rates, and referral patterns within the healthcare system.

Agent Types:

1. General Hospitals: Representing primary healthcare facilities, General Hospitals have a specified success rate for diagnosing patients and a processing time which is equal to 0.6 and 14 (days) respectively. They refer patients to Teaching Hospitals based on a certain condition. The condition requires that biomarker tests are performed for patients with a success rate of 0.6. The biomarker tests that are provided by General Hospital are: PD-L1, ALK, EGFR and KRAS. For each patient, a random number is generated; if this number is less than or equal to the success rate, the patient is diagnosed. Otherwise, the patient is referred to a teaching hospital for further biomarker tests. Additionally, since the test results are prepared after 15 days, the patient will be referred to a teaching hospital after this period.

2. Teaching Hospitals: Teaching hospitals perform the same tests in general hospital (PD-L1, ALK, EGFR and KRAS) in addition to BRAF and ROS1.Teaching Hospital accept patients from General Hospitals and perform the mentioned biomarker tests with the success rate of 0.6. Like the General Hospital, a random number is generated for each patient and if the random number is less than or equal to 0.6, the patient is diagnosed. Otherwise, the patient is referred to an academic hospital. Additionally, since the test results are prepared after 15 days, the patient will be referred to an academic hospital after this period.

3. Academic Hospitals: Academic Hospitals receive patients from Teaching Hospitals and may also receive direct referrals from General Hospital in scenario 2. Like the general

hospital and teaching hospital, biomarker tests are performed for patients. The biomarker tests are the same as the teaching hospital in addition to WGS. Based on the patient and physician willingness for WGS, WGS is performed. The probability of patient and physician willingness for WGS is 0.9. A random number is generated for each patient, if the random number is equal or less than the multiplication of patient and physician willingness, patient goes for the WGS, unless patients go for the same biomarker tests. If the patient goes for the biomarker tests, a random number is generated, if the random number id equal or less than the test success rate which is equal to 0.6 patient is diagnosed, unless patient is not diagnosed.

4. WGS Facility: The Whole Genome Sequencing Facility conducts genomic analyses on referred patients from Academic Hospitals. A random number is generated, if the random number is equal or less than the 0.66 which is the enough tumor percentage for WGS, patient is diagnosed, otherwise patient will receive biomarker tests.

Model Initialization: The model initializes with a predefined number of each type of hospital and the WGS Facility. Each hospital has attributes such as success rates, processing times, patient acceptance rates, and referral probabilities. Agents are placed on a grid representing the physical layout of the healthcare system.

Simulation Process:

1. Patient Flow: Patients enter the healthcare system through General Hospitals, where they may be diagnosed or referred to Teaching Hospitals. Subsequently, patients may be referred to Academic Hospitals for further evaluation or to the WGS Facility for genomic analysis.

2. Diagnosis and Referral: Hospitals success in diagnosing patients is based on random probabilities. Patients who are not diagnosed will be referred to other hospitals or facilities.
3. Data Collection: The model collects data on diagnosed patients, referrals, and total patients for each type of hospital over time.
4. Visualization: After simulation, the model provides visualizations of hospital positions, diagnosed patients over time, and total patients over time.

3.4 Description of the Agent-Based Model Built in Mesa

In this agent-based model, implemented using the Mesa framework, we have structured the model with four distinct agent classes and one overarching model class. Each class serves a specific role within the simulation, contributing to the overall dynamics and interactions within the system.

General Hospital Agent:

- Represents general hospitals that patients can visit.
- Attributes include test success probability rate, patient acceptance, test processing time and cost.
- Behaviors include diagnosing patients and referring them to teaching hospitals if necessary.

```python
#Initializing GeneralHospital class
class GeneralHospital(mesa.Agent):
  def __init__(self, unique_id, model, G_success_rate, G_initial_patient, alpha, G_cost):
    super().__init__(unique_id, model)
    self.agent_type = "GeneralHospital"
    self.success_rate = G_success_rate
    self.initial_patient = G_initial_patient
    self.G_diagnosed = []
    self.closest_teaching_hospital = None
    self.Gprocessing_time = 15
    self.refer_to_T = []
    self.refer_to_A= []
    self.cost = G_cost
    self.G_totalcost = 0
    self.Alpha = alpha
    self.waiting_time = 0  #for all patient in General hospital
    self.referraltime = 0 #for a patient who referred from G to T or A
  def step(self):
    self.G_diagnosed = []
    #perform biomarker tests for patients
    for i in range(1, self.initial_patient + 1):
      Random = np.random.rand()
      if Random <= self.success_rate:
       self.G_diagnosed.append(1)
      else:
        if Random <= self.Alpha:
         self.refer_to_T.append([1,self.model.schedule.steps])
        else:
         self.refer_to_A. append([1, self.model.schedule.steps])

    #calculate total cost for all patients
    self.G_totalcost = self.cost * self.initial_patient
    #calculate waiting time for patients in General Hospital
    self.waiting_time = self.Gprocessing_time * (len(self.G_diagnosed) + len(self.refer_to_A) + len(self.refer_to_T)) + self.referraltime
```

Figure 4: Python code for general hospital agent class.

Figure 4 shows the code for General Hospital agent class. Class General Hospital has been initiated with some parameters such as G_success_rate, G_initial_patient which is the number of patients visiting general hospital by themselves, alpha which is the proportion of sending patients to teaching hospital and G_cost which is the cost of biomarker tests at general hospital. Behaviors such as diagnosing patients and referring patients occur at step function and it will be run at every step. For example, at step function for the general hospital, a random number is generated for patients visiting general hospital, if the random number is less than the test success rate (in this case for general hospital is 0.58), the patient is considered as diagnosed patient, unless the patient will be referred to teaching hospital or academic hospital.

Teaching Hospital Agent:

- Represents general hospitals that patients can visit.
- Attributes include test success probability rate, patient acceptance, test processing time and cost.
- Behaviors include receiving patients from general hospitals, diagnosing patients, and referring them to academic hospitals if necessary.

```
#Initializing TeachingHospital class
class TeachingHospital(mesa.Agent):
  def __init__(self, unique_id, model, T_success_rate, patient_accpt, T_cost):
    super().__init__(unique_id, model)
    self.agent_type = "TeachingHospital"
    self.success_rate = T_success_rate
    self.patient_accpt = patient_accpt
    self.T_diagnosed = []
    self.refer_to_A = []
    self.refer_from_G = []
    self.refer_fromG_history = []
    self.closest_academic_hospital = None
    self.total_p = patient_accpt
    self.Tprocessing_time = 15
    self.cost = T_cost
    self.T_totalcost = 0
    self.T_refercost = 0
    self.waiting_time = 0   #for all patients at Teaching hospital
    self.referraltime = 0   #for a patient who referred from T to A
def step(self):
# Calculate total number of patients in the TeachingHospital
   self.total_p = self.patient_accpt + len(self.refer_from_G)
   self.T_diagnosed = []
   self.refer_fromG_history = len(self.refer_from_G)
   self.refer_from_G = []

  #perform biomarker tests for patients
   for i in range(1, self.total_p + 1 ):
     Random = np.random.rand()
     if Random <= self.success_rate:
       self.T_diagnosed.append(1)
     else:
       self.refer_to_A.append([1,self.model.schedule.steps])

  #calculate total cost for all patients
   self.T_totalcost = self.cost * self.total_p
  #calculate referred cost
   self.T_refercost = self.cost * self.refer_fromG_history
  #calculate waiting time for patients in Teaching Hospital
   self.waiting_time = self.Tprocessing_time * self.total_p + self.referraltime
```

Figure 5: Python code for teaching hospital agent class

Figure 5 shows the code for Teaching Hospital Class. Class Teaching Hospital has been initiated with some parameters such as T_success_rate, patient_accpt which is the number of patients visiting teachingl hospital by themselves, and T_cost which is the cost of biomarker tests at

teaching hospital. Behaviors such as diagnosing patients and referring patients occur at step function and it will be run at every step. For example, at the begging step function for the teaching hospital, the sum of number of patients referred from general hospital and patients who visited the teaching hospital by themselves is calculated, then a random number is generated for total patients, if the random number is less than the test success rate (in this case for teaching hospital is 0.6), the patient is considered as diagnosed patient, unless the patient will be referred to academic hospital.

Academic Hospital Agent:

- Represents academic hospitals that patients can visit.
- Attributes include test success probability rate, patient acceptance, test processing time and cost.
- Behaviors include receiving patients from both general and teaching hospitals for WGS test, diagnosing patients, and referring them to academic hospitals if necessary.

```
#Initializing AcademicHospital class
class AcademicHospital(mesa.Agent):
  def __init__(self, unique_id, model, A_success_rate, A_patient_accpt, dr_lh, p_lh, A_cost):
    super().__init__(unique_id, model)
    self.agent_type = "AcademicHospital"
    self.success_rate = A_success_rate
    self.A_patient_accpt = A_patient_accpt
    self.refer_from_T = []
    self.refer_from_G = []
    self.total_p_a = A_patient_accpt
    self.physician_lh = dr_lh
    self.patient_lh = p_lh
    self.probability = self.patient_lh * self.physician_lh
    self.refer_to_WGS = []
    self.notdiagnosed = []
    self.A_diagnosed = []
    self.soc = []
    self.closest_wgs = None
    self.refer_from_WGS = []
    self.Aprocessing_time = 18
    self.refer_fromT_history = []
    self.refer_fromG_history = []
    self.refer_fromWGS_history = []
    self.cost = A_cost
    self.A_totalcost = 0
    self.A_refercost = 0
    self.waiting_time = 0   #for all patients at Academic hospital
    self.referraltime = 0   #for a patient who referred from A to WGS
def step(self):
  self.total_p_a = self.A_patient_accpt + len(self.refer_from_T) + len(self.refer_from_WGS) + len(self.refer_from_G)
  self.A_diagnosed = []
  self.refer_fromT_history = len(self.refer_from_T)
  self.refer_fromWGS_history = len(self.refer_from_WGS)
  self.refer_fromG_history = len(self.refer_from_G)
  self.refer_from_T = []
  self.refer_from_G = []
  self.refer_from_WGS = []
  self.soc_patients = 0

  # perform biomarker test for patienst
  for i in range(1, self.total_p_a + 1):
    Random = np.random.rand()
    if Random <= self.probability:
      self.refer_to_WGS.append([1, self.unique_id, self.model.schedule.steps])
    else:
      self.soc_patients += 1
      self.soc.append([1,self.model.schedule.steps])
      Random3 = np.random.rand()
      if Random3 <= self.success_rate:
        self.A_diagnosed.append(1)
      else:
        self.refer_to_WGS.append([1, self.unique_id, self.model.schedule.steps])
  #calculate total cost for all patients
  self.A_totalcost = self.cost * self.soc_patients
  #calculate referred cost
  self.A_refercost = self.cost * self.refer_fromT_history +  self.cost * self.refer_fromWGS_history
  #calculate waiting time for all patients in Academic Hospital
  self.waiting_time = self.Aprocessing_time * self.total_p_a + self.referraltime
```

Figure 6: Python code for academic hospital agent class.

Figure 6 shows the code for Academic Hospital Agent class. Class Academic Hospital has been initiated with some parameters such as A_success_rate, A_patient_accpt (which is the number of patients visiting teachingl hospital by themselves) , dr_lh, p_lh which are docotr's and patients likelihood for performing WGS test. s, and A_cost which is the cost of biomarker tests at academic hospitals. Behaviors such as diagnosing patients and referring patients occur at step function and it will be run at every step. For example, at the begging step function for the academic hospital, the sum of number of patients referred from general hospital and teaching hospital and patients who visited the academic hospital by themselves is calculated, then a random number is generated for total patients, if the random number is less than the test success rate (in this case for teaching hospital is 0.6), the patient is considered as diagnosed patient, unless the patient will be referred to WGS facility for WGS test.

WGS Facility Agent:

- Represents WGS facility that patients will be referred to.
- Attributes include test success probability rate, capacity limitation, test processing time and cost.
- Behaviors include receiving patients from academic hospital for WGS test, diagnosing patients, and referring back them to academic hospitals if necessary**.**

```python
#Initializing WGSFacility class
class WGSFacility(mesa.Agent):
  def __init__(self, unique_id, model, percentage, WGS_cost, capacity):
    super().__init__(unique_id, model)
    self.agent_type = "WGSFacility"
    self.refer_from_A = []
    self.tumor_percentage = percentage
    self.wgs_diagnosed = []
    self.refer_to_A = []
    self.total_wgs = 0
    self.WGSprocessing_time = 14
    self.refer_fromA_history = []
    self.cost = WGS_cost
    self.WGS_totalcost = []
    self.WGS_refercost = 0
    self.waiting_time = 0
    self.referraltime = 0
    self.in_queue = []
    self.capacity = capacity
    self.wgs_patients = 0
  def step(self):
    self.total_wgs = len(self.refer_from_A) + len(self.in_queue)
    wgs_patients = self.refer_from_A
    self.in_queue = []
    self.wgs_diagnosed = []
    #self.refer_fromA_history= len((self.refer_from_A))
    self.refer_from_A = []
    remaining_capacity = self.capacity

    for i in range(1, self.total_wgs + 1):
      if remaining_capacity == 0:
        self.in_queue.append(1)
      else:
        remaining_capacity -= 1
        Random4 = np.random.rand()
        if Random4 <= self.tumor_percentage:
         self.wgs_diagnosed.append(1)
        else:
         if i <= len(wgs_patients):
          self.refer_to_A.append([1, wgs_patients[i-1][1], self.model.schedule.steps])

    #calculate total cost for all patients
    self.WGS_totalcost = self.cost * self.total_wgs
    #calculate referred patients cost
    self.WGS_refercost = self.cost * self.total_wgs
    #calculate waiting time for all patients in Academic Hospital
    self.waiting_time =  self.WGSprocessing_time *  self.refer_fromA_history
```

Figure 7: Python code for academic hospital agent class.

Figure 7 shows the code WGS Facility Agent class. Class WGS Facility has been initiated with some parameters such as percentage (which is the tumor percentage), WGS_cost, and capacity which is the capacity limitation for WGS facility. Behaviors such as diagnosing patients and referring patients occur at step function and it will be run at every step. For example, at the begging step function for the WGS facility, the sum of number of patients referred from academic hospital and patients who are in the queue (due to the capacity limitation a queue will be produced) is

calculated, then a random number is generated for total patients, if the random number is less than the percentage (in this case wgs facility is 0.66), the patient is considered as diagnosed patient, unless the patient will be referred back to academic hospital for SOC.

Model Class

The Hospital Model class is a simulation model for different types of hospitals, created using the Mesa framework. Some values and referral logic, such as finding the closest distance between hospitals or the closest distance for patients to the hospitals, have been initialized in the model class.:

```
for i in range(1, self.num_general + 1):
    G_hospital = GeneralHospital(i, self, G_success_rate = 0.58, G_initial_patient = 2, alpha = self.alpha_value, G_cost = 334.14)
    self.schedule.add(G_hospital)
    # Add the agent to a random grid cell
    x = random.uniform(0, self.space.x_max)
    y = random.uniform(0, self.space.y_max)
    # place_agent method adds the coordinate to the agent automatically
    self.space.place_agent(G_hospital, (x,y))
```

Figure 8: General Hospital agent initialization

For example, in Figure 8, we have initialized the General Hospital agents with their values. The code for Teaching Hospitals, Academic Hospitals, and WGS facilities is the same as for General Hospitals, with different values and parameters. However, the structure of the code and the logic is the same.

```
#Calculate the closest teaching hospital for all general hospitals
for agent in self.schedule.agents:
    if agent.agent_type == "GeneralHospital":
      closest_teaching_hospital = None
      closest_academic_hospital = None
      closest_distanceT = float('inf')
      closest_distanceA = float('inf')  # Initialize with a large value
      for th in self.schedule.agents:
         if th.agent_type == "TeachingHospital":
             distance = self.calculate_distance(agent.pos, th.pos)
             if distance < closest_distanceT:
                 closest_distanceT = distance
                 closest_teaching_hospital = th.unique_id
      agent.closest_teaching_hospital = closest_teaching_hospital

      for ac in self.schedule.agents:
         if ac.agent_type == "AcademicHospital":
             distance = self.calculate_distance(agent.pos, ac.pos)
             if distance <  closest_distanceA:
                 closest_distanceA = distance
                 closest_academic_hospital = ac.unique_id
      agent.closest_academic_hospital = closest_academic_hospital
```

Figure 9: Code for finding the closest distance for general hospital from teaching and academic hospitals.

In figure 9, we have defined how to find the closest teaching and academic hospitals for general hospitals.

```
#define the referral logic
  def step(self):
        for agent in self.schedule.agents:
            if agent.agent_type == "GeneralHospital":
              for th in self.schedule.agents:
                if th.agent_type == "TeachingHospital" and th.unique_id == agent.closest_teaching_hospital:
                  refer_to_T  = [r for r in agent.refer_to_T if self.schedule.steps-r[1]>= agent.Gprocessing_time ]
                  th.refer_from_G.extend(refer_to_T)
                  agent.refer_to_T = [r for r in agent.refer_to_T if self.schedule.steps-r[1]< agent.Gprocessing_time ]


              for ac in self.schedule.agents:
                if ac.agent_type == "AcademicHospital" and ac.unique_id == agent.closest_academic_hospital:
                  refer_to_A = [r for r in agent.refer_to_A if self.schedule.steps-r[1] >= agent.Gprocessing_time ]
                  ac.refer_from_G.extend(refer_to_A)
                  agent.refer_to_A = [r for r in agent.refer_to_A if self.schedule.steps-r[1] < agent.Gprocessing_time ]
```

Figure 10: Code for referral logic from general hospital to teaching or academic hospital.

In figure 10, the referral logic for general hospital has been defined.

3.5 Results

3.5.1 Base Scenario Result

In the base scenario, a patient who is not diagnosed at the general hospital will be referred to a teaching hospital. Results of the base scenario is as below:

For both scenarios, the simulation model has been run for various types of hospitals, including 6 general hospitals, 3 teaching hospitals, 1 academic hospital, and 1 WGS facility, over 200-time steps. Each time step represents one day and simulates the visits of 2 patients to general hospitals, and 1 patient each to teaching and academic hospitals. The total annual patients are 5313, with 60% visiting general hospitals, 29% visiting teaching hospitals, and 11% visiting academic hospitals [1].

To calculate the total cost for general hospitals, we multiply the number of patients who visit the general hospital by $334.14, which is the cost of performing a biomarker test at a general hospital. For example, if there are 2 patients at the general hospital, the total cost would be 2 multiplied by 334.14.

$C_i = Cost\ of\ each\ Hospital \qquad i \in I$

$P_i = Cost\ of\ Biomarker\ Tests\ per\ Patient \quad i \in I$

$N_i = Number\ of\ Total\ Patients\ at\ each\ hospital \qquad i \in I$

$C_i = P_i N_i$

Example:

Number of total patients at the general hospital (N_i): 2

Cost per patient (P_i): 334.14

Total cost (C_i) = 2 * 334.13 = 668.28

To calculate the total cost for teaching hospitals, we multiply the number of patients who visit the teaching hospital, plus the number of patients who are referred to the teaching hospital, by $536.02. To calculate the total cost for an academic hospital, we base the calculation on whether patients undergo SOC or WGS. If patients undergo SOC, we multiply the number of patients by $536.02. If patients undergo WGS, we multiply the number of patients by $2925.25; however, this cost will be considered the total cost for the WGS facility, not for the academic hospital.

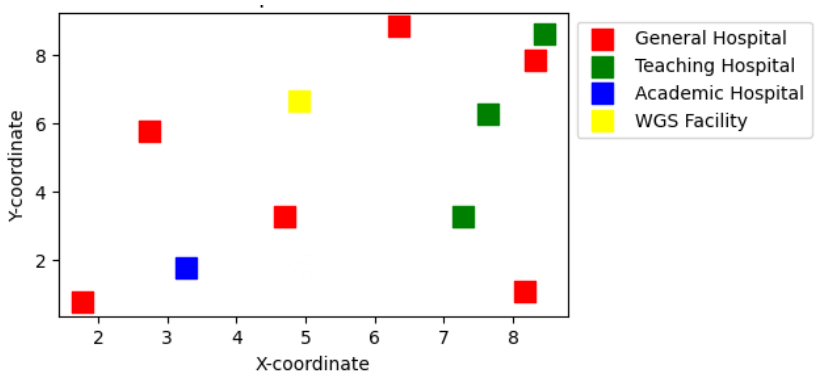

Figure 11: Map of hospital positions.

Figure 11 displays the randomly generated locations of various hospitals. The red markers indicate general hospitals, located at six different sites. The green markers represent teaching hospitals, located at three distinct locations. The blue marker shows the location of an academic hospital, while the yellow marker represents the WGS facility, each at a single location.

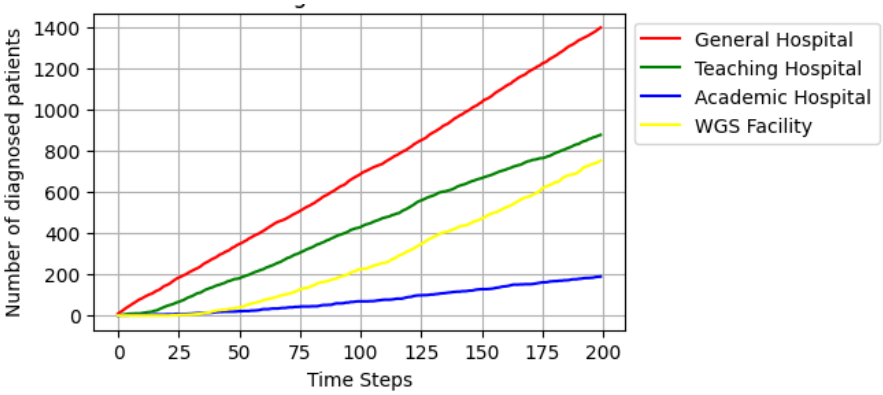

Figure 12: Trends in patients diagnosed at different types of hospitals.

Figure 12 presents the cumulative number of diagnosed patients at different types of hospitals over a span of 200-time steps. The time step unit in this simulation is considered to be a day. This means that 200-time steps are equal to 200 days. The red line tracks diagnosed patients at general hospitals. This line shows an almost linear increase in the number of diagnosed patients because patient's arrival has a uniform distribution and number of diagnosed patients reaching up to approximately 1400 by the 200th time step. The green line tracks diagnosed patients at teaching hospitals. This line also demonstrates a linear increase, though at a slower rate compared to the general hospitals because there are fewer patients in teaching hospitals. The line culminates at around 900 diagnosed patients. The yellow line indicates the cumulative diagnoses at the WGS facility and reaching nearly 700 patients by the end of the observed period. The number of diagnosed patients at the WGS facility is near zero at the beginning of the period because when patients visit academic hospitals or are referred there from teaching hospitals, it takes 14 days to process the WGS test results, in addition to the waiting time for results at general or teaching hospitals. Therefore, in the initial steps of the simulation, there are no patients at the WGS facility. The blue line represents the academic hospital, which has the lowest number of diagnosed patients the four, ending with around 200 diagnosed patients by the 200th time step.

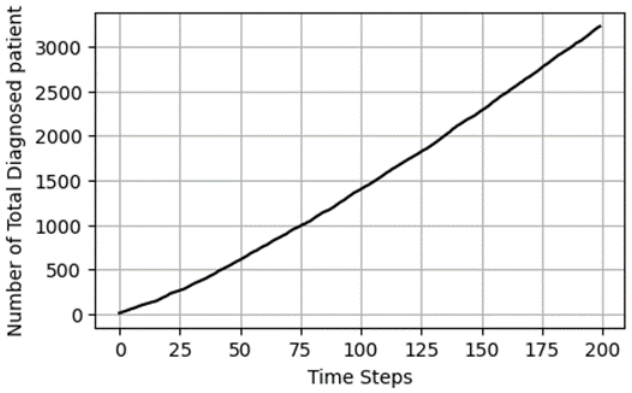

Figure 13: Trends in number of diagnosed patients for all hospitals.

Figure 13 illustrates the cumulative total number of diagnosed patients over a series of time steps. The graph displays a single black line showing a continuous and steady increase in the total number of diagnosed patients over time. The reaches approximately 3200 diagnosed patients by the 200th time step. The total number of diagnosed patients in the base scenario is 3229.

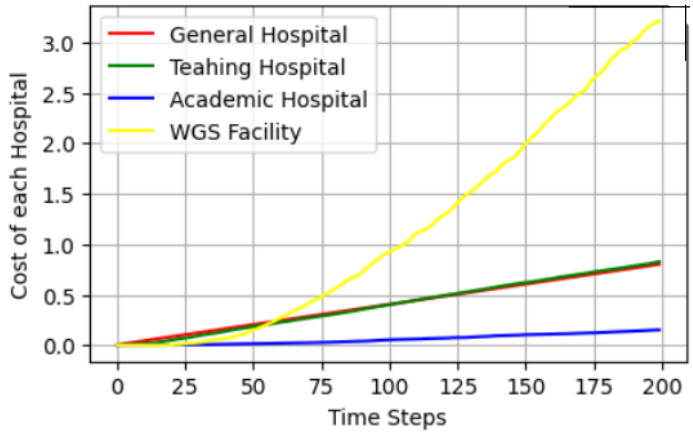

Figure 14: Comparative analysis of hospital costs.

Figure 14 presents the cumulative costs for four distinct healthcare facilities over a sequence of 200-time steps. The red line represents the general hospital, the green line is for the teaching hospital, the blue line shows the academic hospital, and the yellow line corresponds to the WGS facility. The lines track the total cost associated with each facility, with the WGS facility seeing a notable surge in costs as time progresses. The WGS facility has the highest cumulative cost, increasing steeply and significantly over time, indicating the high expenses involved in whole genome sequencing. WGS test cost is 2925.25 euros per patient while the biomarker tests provided by general, teaching, and academic hospital is 334.14 euros, 536.02 euros and 536.02 euros.

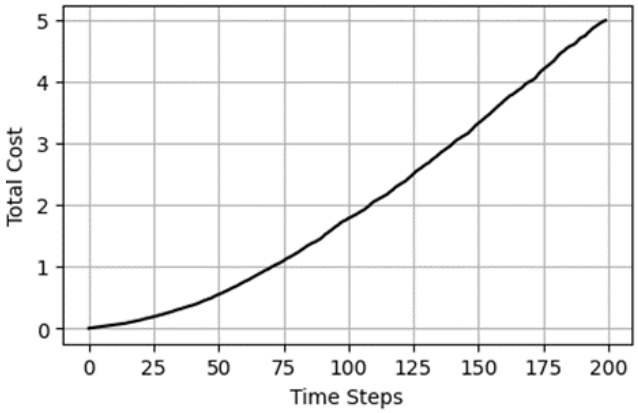

Figure 15: Trends in total cost for all hospitals in base scenario

Figure 15 shows the cumulative cost for all hospital types combined over 200 time steps.

3.5.1.1 Result with WGS Capacity in base scenario

The base scenario model has been run again with WGS capacity. WGS capacity is 1600 biopsies per year [1]. In this model (base scenario), there is one WGS facility and the capacity has been considered 4 per day.

(a) (b)

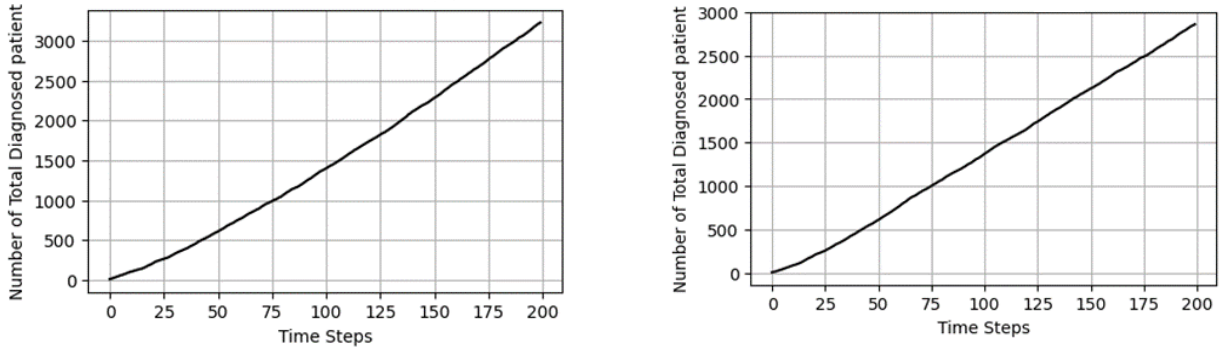

Figure 16: Trends in number of diagnosed patients for all hospitals without WGS capacity (a) and with WGS capacity (b)

Figures 16 depict the cumulative number of diagnosed patients across all hospitals over 200-time steps. The key distinction between these two figures lies in the WGS capacity. Figure 16 (a) is without WGS capacity, and figure 16 (b) is with WGS capacity. When the capacity of the WGS facility is accounted for, there is a reduction in the total number of diagnosed patients. The number

of total diagnosed patients with WGS capacity is equal to 2859 patients over 200-time steps. The number of total diagnosed patients without WGS capacity is equal to 3229 patients over 200-time steps.

(a) (b)

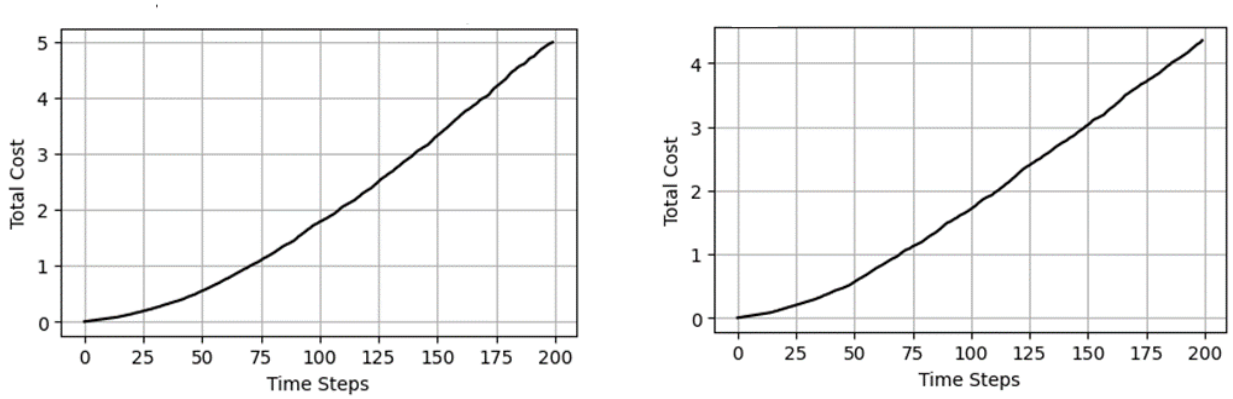

Figure 17: Trends in cumulative cost for all hospitals without WGS capacity (a) and with WGS capacity (b)

Figures 17 illustrates the cumulative total cost for all hospitals with and without WGS capacity. Comparing these two graphs reveals that the total cost is higher when WGS capacity is not implemented, figure 17 (a). Specifically, the total cost with WGS capacity is 4,356,415.97, while the total cost without WGS capacity is 5,200,000.

3.5.2 Scenario 2 Result

In scenario 2, if a patient is not diagnosed at the general hospital, they will be referred directly to the academic hospital. Patients are no longer referred to the teaching hospital from the general hospital. Like the base scenario, the simulation model has been run for various types of hospitals including 6 general hospitals, 3 teaching hospitals, 1 academic hospital and 1 WGS facility running through 200 steps. Each step simulates the visits of 2 patients to general hospitals, and 1 patient each to teaching and academic hospitals. The total annual patients are 5313 which 60% visit general hospital, 29% visit teaching hospital and 11% visit academic hospital.[1]

(a) (b)

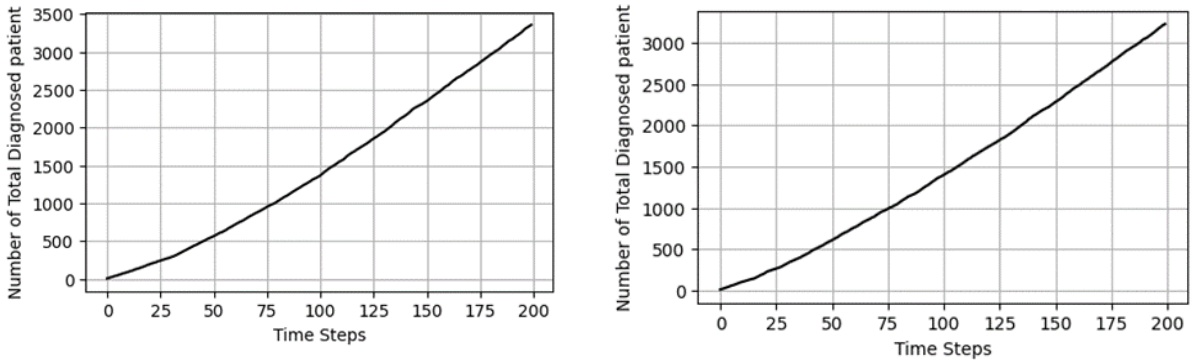

Figure 18: Cumulative diagnosed patients for all hospitals - scenario 2 (a) and base scenario (b)

Figure 18 illustrates the cumulative total number of diagnosed patients over a series of time steps. Figure 18 (a) shows the cumulative diagnosed patients in scenario 2 which the number of diagnosed patients is more than the base scenario. The total number of diagnosed patients in scenario 2 is 3356 which is more than the base scenario which is equal to 3229. By analyzing this graph, we realized that when capacity is not applied for WGS facility, it is better to send patients to academic hospital directly from general hospital.

(a) (b)

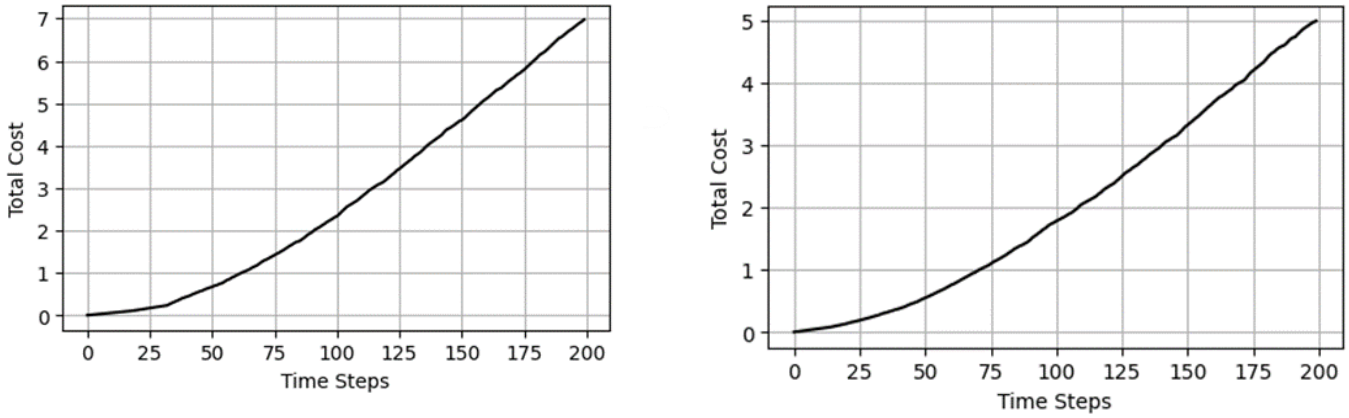

Figure 19:Trends in cumulative cost for all hospitals - scenario 2 (a) and base scenario (b)

Figure 19 shows the cumulative cost for all hospital types combined over 200 time steps. Total cost for scenario 2 is more than total cost for the base scenario.

3.5.2.1 Result with WGS capacity limitation in scenario 2

The scenario 2 model has been run again with WGS capacity. WGS capacity is 1600 biopsies per year.[1] In this model there is one WGS facility and the capacity limit has been considered 4 per day.

(a) (b)

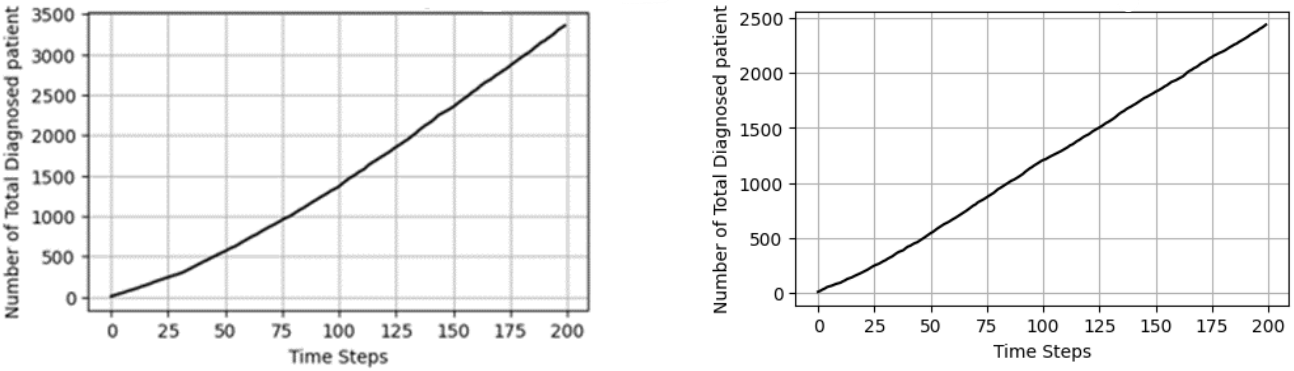

Figure 20: Trends in cumulative diagnosed patient for all hospitals in scenario 2 without WGS capacity (a) and with WGS capacity (b)

Figures 20 depict the count of diagnosed patients across all hospitals over 200-time steps. The key distinction between these two figures lies in the WGS Capacity. When the capacity of the WGS facility is accounted for, there is a reduction in the total number of diagnosed patients.

(a) (b)

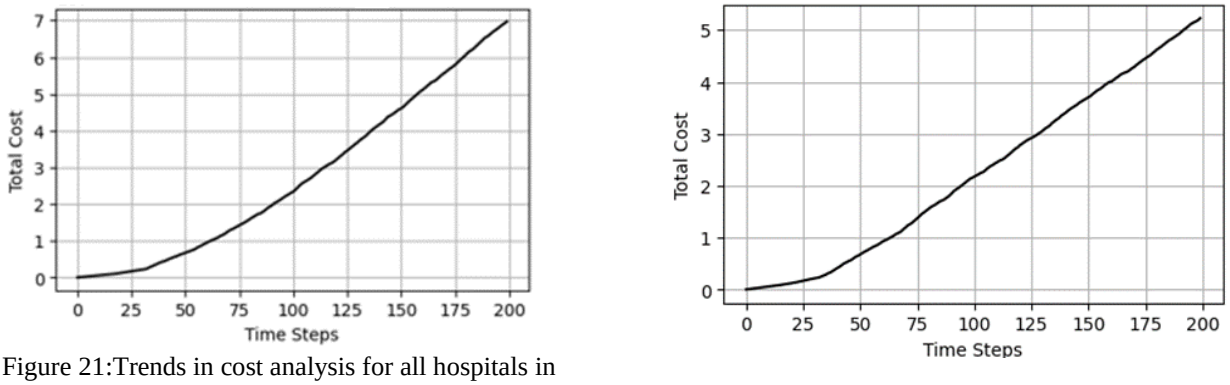

Figure 21:Trends in cost analysis for all hospitals in scenario 2 without WGS capacity (a) and with WGS capacity (b)

Figures 21 illustrates the total cost for all hospitals without and with WGS capacity. Comparing these two graphs reveals that the total cost is higher when WGS capacity is not implemented.

Specifically, the total cost with WGS capacity is 5,230,000 while the total cost without WGS capacity is 6,970,000.

3.5.3 Comparative analysis for number of diagnosed patients

Here is a table which shows the number of diagnosed patients in different scenarios with and without WGS capacity.

Table 3: Comparison of Total diagnosed patients for Hospitals with and without WGS Capacity under Different Scenarios.

	Without WGS Capacity	With WGS Capacity
Base Scenario	3229	2859
Scenario 2	3356	2437

As shown in table 2, sending patients to WGS facility sooner can increase the number of diagnosed patients but sending everybody is not optimal as the capacity limit may cause people to wait even longer.

3.5.4 Comparative analysis for total cost

Table 3 shows total cost in different scenarios with and without WGS capacity.

Table 4: Comparison of Total cost for Hospitals with and without WGS Capacity under Different Scenarios.

	Without Capacity	Cost per diagnosed	With Capacity	Cost per diagnosed
Base Scenario	5,200,000	1610	4,356,415.97	1523
Scenario 2	6,970,000	2076	5,230,000	2146

As shown in table 3, the cost is still higher if we send everybody to WGS. Also implementing WGS capacity results in a decrease in total cost. This is because some patients are placed in a queue, causing delays before they can proceed to the WGS test.

3.5.5 Comparing Scenarios without WGS Capacity

We compared the base scenario and scenario 2 in terms of total diagnosed patients, total cost, and the ratio of total diagnosed patients to total cost, as shown in three graphs. In this comparison, we considered no WGS capacity for the WGS facility. We defined a parameter named alpha. Alpha is a parameter that specifies the proportion of patients being referred to teaching or academic hospitals. When alpha equals 1, patients follow the base scenarios. When alpha equals 0, all patients are directly referred to academic hospitals from general hospitals.

3.5.5.1 Comparing diagnosed patients for two scenarios

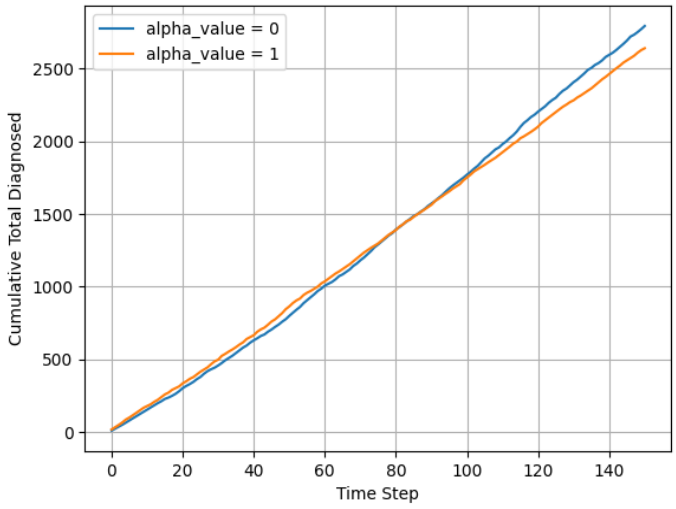

Figure 22: Total diagnosed patients for all hospitals.

Figure 22 shows the cumulative number of diagnosed patients over time for two different alpha values (0 and 1). When alpha equals 1, patients follow the base scenarios which is sending patients to teaching hospital from general hospital. When alpha equals 0, all patients are directly referred

to academic hospitals from general hospitals. The graph indicates that the cumulative number of diagnosed patients is higher when patients are referred to academic hospital from general hospital. This suggests that the parameter controlled by the alpha value influences the rate at which patients are diagnosed, with alpha_value = 0 resulting in a higher cumulative total diagnosed patient. For this graph, we have considered the time steps from step 50.

3.5.5.2 Comparing total cost for two scenarios

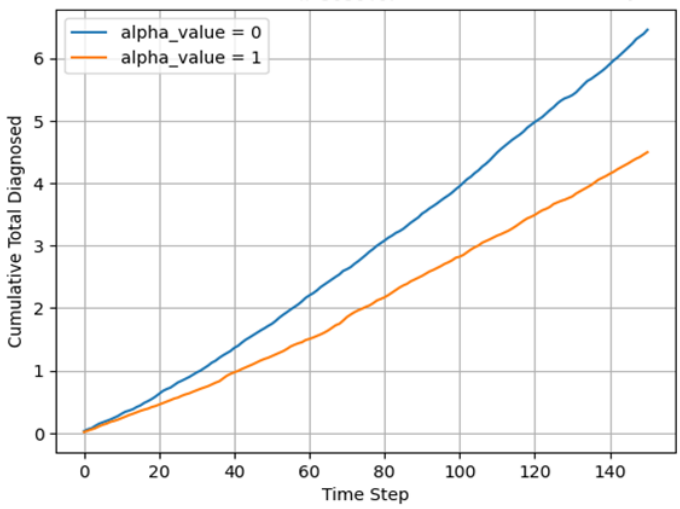

Figure 23: Comparative analysis of cost for two scenarios.

Figure 23 shows the cumulative total cost over time for two different alpha values (0 and 1). The graph indicates that the cumulative total cost is higher when the alpha value is 0 (direct referral from general hospital to academic hospital) compared to when it is 1(referral from general hospital to academic hospital). This suggests that the parameter controlled by the alpha value ,which is the direct referral from general hospital to academic hospital (scenario 2), influences the total cost, with alpha_value = 0 resulting in a higher cumulative total cost. For this graph, we have considered the time steps from step 50.

3.5.5.3 Comparing a ratio for two scenarios

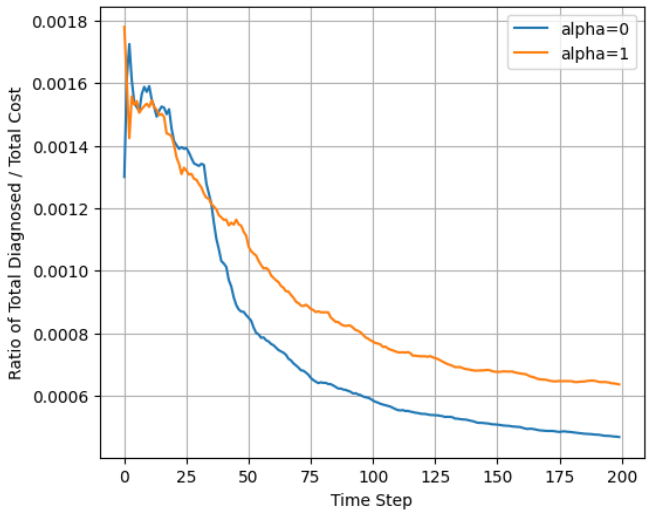

Figure 24: Comparative ratios for two scenarios

Figure 24 displays the ratio of the cumulative number of diagnosed patients to the cumulative total cost over time for two different scenarios: alpha = 0 and alpha = 1. The results indicate that alpha = 1 maintains a higher efficiency in terms of diagnosis per unit cost compared to alpha = 0. This finding highlights the importance of the alpha parameter in optimizing the cost-effectiveness of diagnostic strategies in the simulation. Understanding this impact is crucial for making informed decisions on resource allocation and improving the overall efficiency of healthcare interventions.

3.5.6 Comparing Scenarios with WGS Capacity

We compared the base scenario and scenario 2 in terms of total diagnosed patients, total cost, and the ratio of total diagnosed patients to total cost, as shown in three graphs. In this comparison, we considered a WGS capacity for the WGS facility which is 4 per day.

3.5.6.1 Comparing diagnosed patients for two scenarios

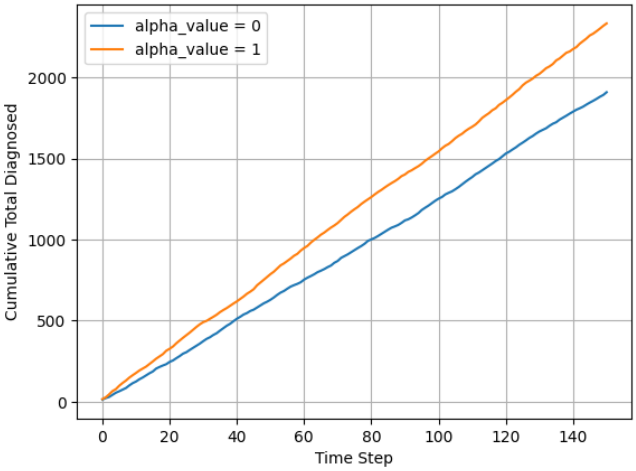

Figure 25: Comparative analysis of diagnosed patients for two scenarios.

Figure 25 shows the cumulative number of diagnosed patients over time for two different alpha values (0 and 1). The graph indicates that the cumulative number of diagnosed patients is higher when the alpha value is 1 compared to when it is 0. This suggests that the base scenario has a greater number of diagnosed patients when WGS capacity is applied. For this graph, we have considered the time steps from step 50.

3.5.6.2 Comparing total cost for two scenarios

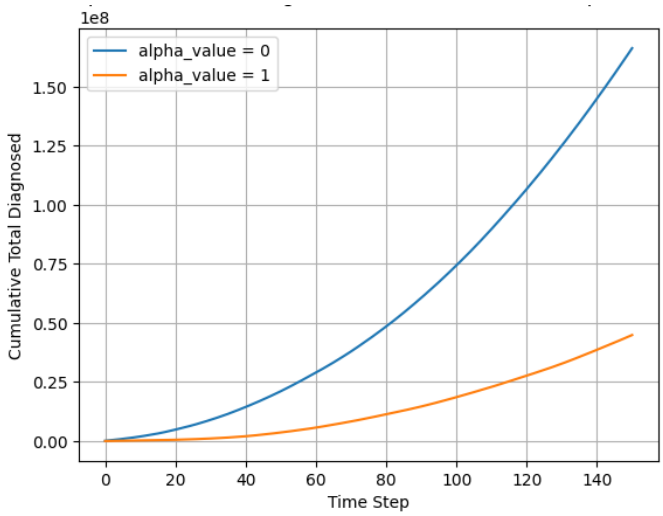

Figure 26: Comparative analysis of total cost for two scenarios.

Figure 26 shows the cumulative total cost over time for two different alpha values (0 and 1). The graph indicates that the cumulative total cost is higher when the alpha value is 0 compared to when it is 1. This suggests that the parameter controlled by the alpha value ,which is the direct referral from general hospital to academic hospital (scenario 2), influences the total cost, with alpha_value = 0 resulting in a higher cumulative total cost. For this graph, we have considered the time steps from step 50.

3.3.6.3 Comparing a ratio for two scenarios

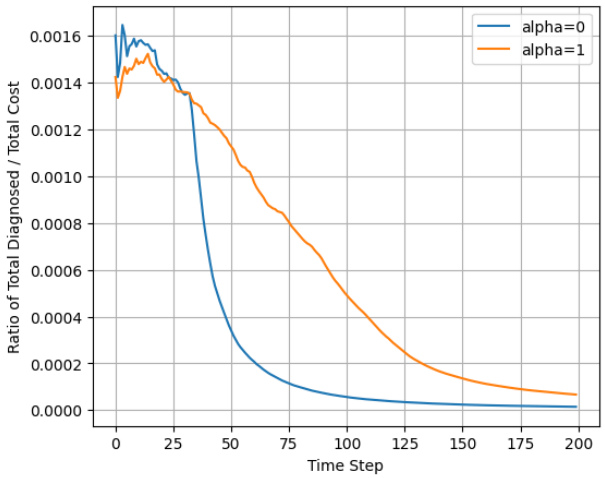

Figure 27: Comparative analysis of ratios for two scenarios.

Figure 27 displays the ratio of the cumulative number of diagnosed patients to the cumulative total cost over time for two different scenarios: alpha = 0 and alpha = 1. The graph provides a clear comparison of the ratio of total diagnosed patients to total cost for two different alpha values. The results indicate that alpha = 1 maintains a higher efficiency in terms of diagnosis per unit cost compared to alpha = 0.

3.5.7 Sensitivity Analysis on WGS Capacity

Alpha is a parameter that specifies the proportion of patients being referred to teaching or academic hospitals. When alpha equals 1, patients follow the base scenarios. When alpha equals 0, all patients are directly referred to academic hospitals from general hospitals. After analyzing the base

scenario and scenario 2 with and without WGS capacity, we observed that applying WGS capacity to the model results in a reduction in the number of diagnosed patients. Additionally, the base scenario has more diagnosed patients when WGS capacity is implemented. In contrast, without WGS capacity, scenario 2 has more diagnosed patients. Consequently, we conducted a sensitivity analysis on WGS capacity. We conducted a sensitivity analysis for different WGS capacities set at 4, 8, 16, and 32.

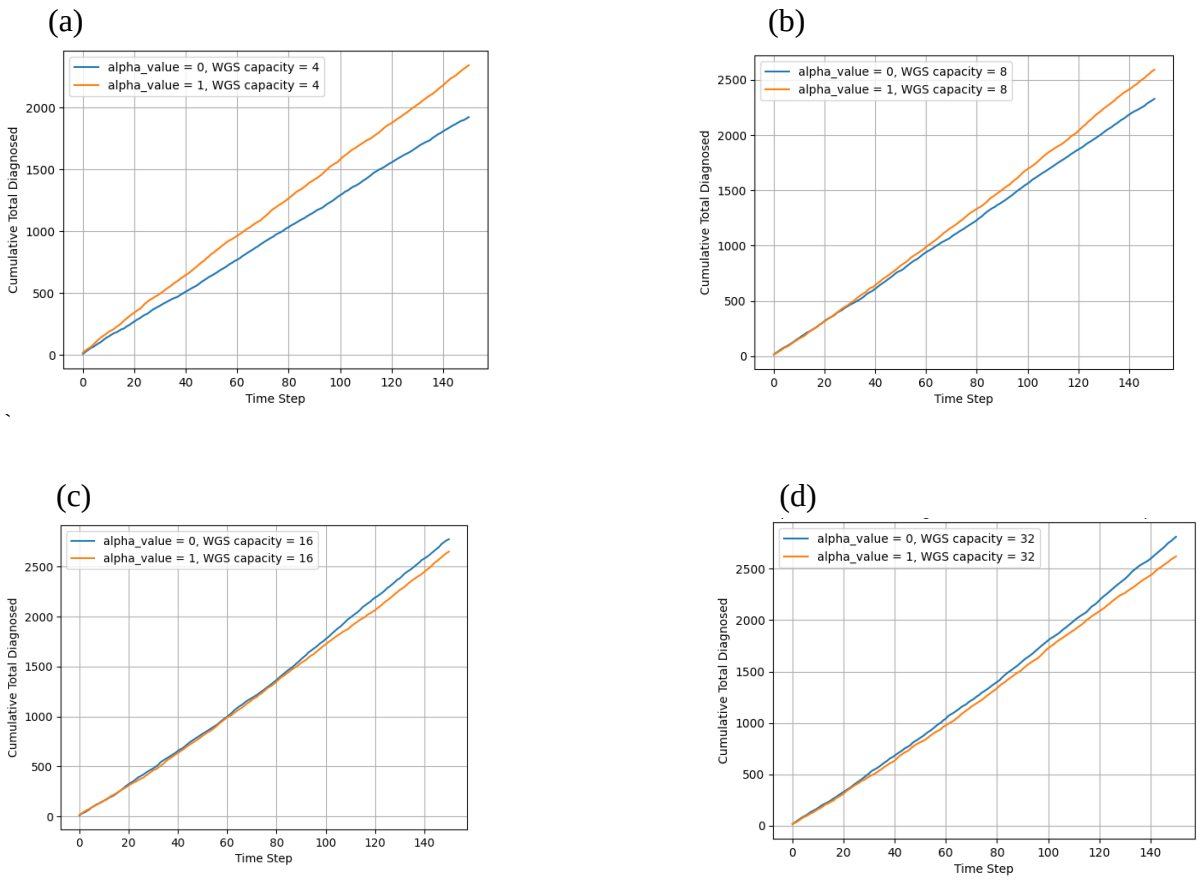

Figure 28: Comparative analysis of diagnosed patients.

Table 5: Comparison of Total diagnosed patients for Hospitals with and without WGS Capacity under Different Scenarios.

WGS Capacity	4	8	16	32
Base Scenario	2859	3139	3200	3264
Scenario 2	2437	2905	3328	3305

Here is the table comparing different WGS capacities. The difference between a WGS capacity of 32 and unlimited capacity is minimal, so it may be more practical to set the WGS capacity to 32 instead of unlimited.

3.6 Conclusion

ABS offers a robust and dynamic framework for addressing the complexities inherent in healthcare systems, specifically when analyzing the intricate interactions between hospitals. By conceptualizing hospitals as autonomous agents with distinct behaviors and interaction patterns. Within this ABS framework, our analysis has highlighted that Scenario 2 sees a higher number of patients successfully diagnosed when the WGS capacity is 32 per day.

CHAPTER IV: AGENT BASED MODEL WITH CONSIDERING PATIENTS AS AGENTS

In Chapter IV, the model is extended to include patients as agents, introducing a new agent class named 'Patient'. The reason that we considered a different agent class for patient is that it allows us to track each patient's diagnostic pathway individually and considers the possibility of patient death due to prolonged waiting times for diagnosis. Additionally, it provides a more accurate calculation of the total cost by accounting for waiting times. Despite this significant change, the behavior of the hospital and other parameters—such as the cost of biomarker tests, processing time, and WGS capacity—remained unchanged. The primary difference between the models in Chapters III and IV lies in the inclusion of patients as active agents in the simulation, providing a more comprehensive analysis of the system's behavior.

4.1 Model's Assumptions

In the Chapter IV model, we have made a few changes, such as considering patients as an agent, defining a death rate. Other assumptions remain the same as in the model in Chapter III. Here are the agents with their attributes:

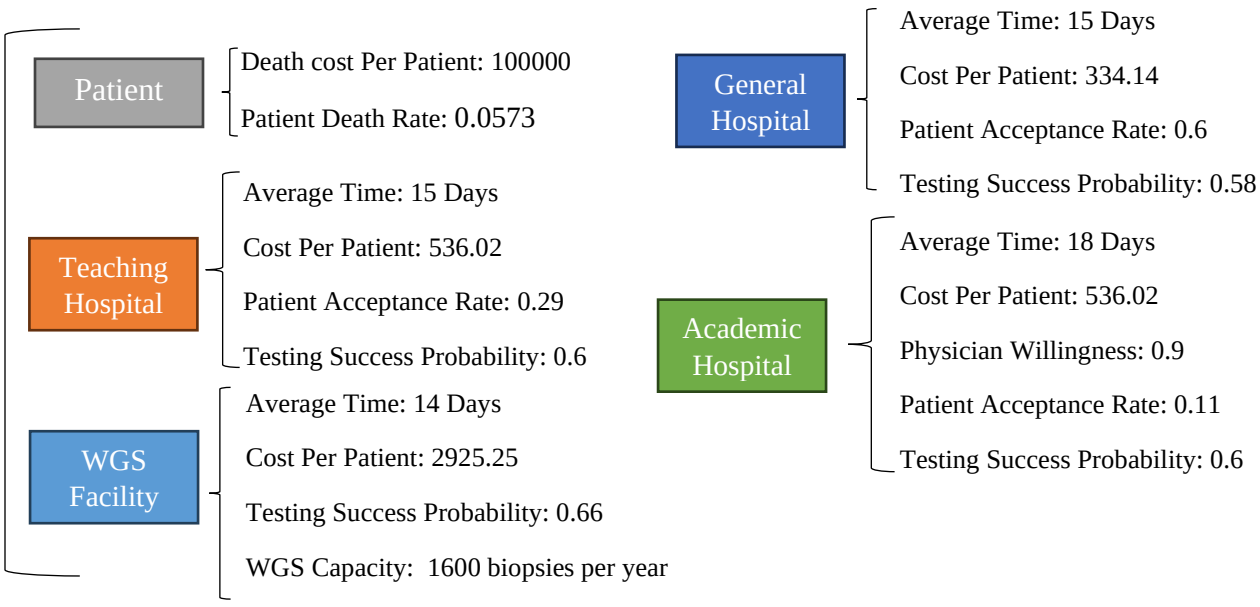

Figure 29: Agent's Attributes

4.1.1 Agent-Based Patient Distribution

In this model, we have considered patients as an agent class. For each patient, a random location is generated in a continuous space, and they will visit general, teaching, or academic hospitals based on their proximity. Annually, the expected number of patients is 5,313. Given this, we have modeled 3,400 patients over 200-time steps, with 60% visiting General Hospitals, 29% visiting Teaching Hospitals, and 11% visiting the Academic Hospital [1].

4.1.2 Patient Mortality Considerations in Stage IV NSCLC

In the latest ABS model, which includes patients as agents, we also account for the probability of patients dying before receiving a diagnosis. Since the patients are in stage IV NSCLC, they are particularly vulnerable. At each step, the model checks if the patients are alive or dead. If a patient is dead, they do not proceed to the next step in the model and a huge cost is added to the patient's cost. We should mention this point that the results for comparing scenarios are without and with death probability. It means that one time we do not consider a death rate probability for patients, and we run the model with this assumption that patients are alive in the simulation, and they do not die, and one more time we assume that patients die in the simulation and we define a probability for death which we explain more in the section which is the results with death rate probability.

4.1.2.1 Survival Rate in Non-small-cell Lung Cancer

Survival rates for stage 4 lung cancer vary based on the type of lung cancer, either non-small cell lung cancer or small cell lung cancer. In this research, we are focusing on non-small cell lung cancer (NSCLC). The extent of its spread and the individual's overall health also plays a significant role. According to the American Cancer Society, NSCLC is the most prevalent type, accounting for about 80 to 85 percent of cases. Their data indicates that the five-year survival rate for

individuals with stage 4 NSCLC, where the cancer has spread to distant parts of the body, is 9 percent [26].

4.1.2.2 Calculating the Death Rate from Survival Rate

The survival rate provided is for a five-year period. Specifically, for individuals with stage 4 non-small cell lung cancer (NSCLC), the five-year survival rate is 9%. This means that 9% of individuals diagnosed with stage 4 NSCLC are expected to survive for at least five years after their diagnosis.

To calculate the death rate, we need to determine the proportion of individuals who do not survive the five-year period. The death rate is essentially the complement of the survival rate. The five-year survival rate (SR) is 9%. The death rate is the percentage of individuals who do not survive for five years. Here's how we can calculate it:

Death Rate = 100% - Survival Rate

Death Rate = 100% - 9%

Death Rate = 91%

Therefore, the five-year death rate for individuals with stage 4 NSCLC is 91%. To accurately model the progression of stage 4 non-small cell lung cancer (NSCLC) on a daily basis, it is essential to convert the five-year survival rate into a daily survival rate. The five-year survival rate, which is 9% for stage 4 NSCLC, represents the proportion of patients expected to survive five years post-diagnosis. This rate must be translated into a daily probability to be used in a day-by-day simulation model. The conversion involves calculating the daily survival rate (DSR) by taking the 1825th root of the five-year survival rate, as there are 1825 days in five years (5 years × 365 days/year). Mathematically, this is expressed as DSR = (5Y SR)^(1/1825). For a five-year survival

rate of 0.09, the DSR is approximately 0.999427, indicating that each day, a patient has a 99.9427% chance of surviving. Consequently, the daily death rate (DDR) is 1 - DSR, or approximately 0.0573%. This daily rate allows for a realistic and granular simulation of patient outcomes over time, aligning with the long-term survival statistics.

Five years: 5 years × 365 days = 1825 days

Daily Survival Rate (DSR): $(DSR)^{1825}$ = 5Y SR

$DSR = 0.09^{1/1825} = 0.999427$

The daily survival rate (DSR) is approximately 0.999427. This means that each day, a patient has a 99.9427% chance of surviving. Conversely, the daily death rate (DDR) is:

$DDR = 1-DSR \approx 1 - 0.999427 = 0.000573$ or 0.0573%

4.2 Comparing scenarios with WGS Capacity & Death Rate

We have included 3,400 patients in our model. The distribution of facilities consists of 6 General Hospitals, 3 Teaching Hospitals, 1 Academic Hospital, and 1 WGS facility. Annually, the expected number of patients is 5,313. Given this, we have modeled 3,400 patients over 200-time steps, with 60% visiting General Hospitals, 29% visiting Teaching Hospitals, and 11% visiting the Academic Hospital. Since patients are considered as agents in our model, each patient is assigned a random location. Both hospital and patient locations are generated randomly within a continuous space, and patients choose hospitals based on proximity [1]. We have considered a death rate for patients, but we do not have include a cost for death and waiting time in this section. As explained in Section 4.1.2.2, the daily death rate (DDR) is 0.0573.

In the base scenario, patients are referred from General Hospital to Teaching Hospitals, and from Teaching Hospitals to Academic Hospitals if they have not been diagnosed. We have introduced a parameter called alpha. When alpha is equal to 1, it indicates that patients are referred to Teaching Hospitals from General Hospitals and when it is equal to 0, it indicates that patients are referred to Academic Hospitals from General Hospitals. Also, the WGS capacity is equal to 4 per day.

$\alpha = 1 \longrightarrow$ Refer patients to Teaching Hospital (Base Scenario)

$\alpha = 0 \longrightarrow$ Refer patients to Academic Hospital (Scenario 2)

4.2.1 Hospital and Patient Position

In this chapter, we have considered patients as agents and here is the code for initializing patient agent class:

```
class Patient(mesa.Agent):
  def __init__(self, unique_id, model):
    super().__init__(unique_id, model)
    self.agent_type = "Patient"
    self.closest_hospital_type = None
    self.closest_hospital_unique_id = 0
    self.entry_time = 0
    self.diagnosed_time = 0
    self.refer_to_T_time = 0
    self.refer_to_A_time = 0
    self.refer_to_WGS_time = 0
    self.patient_cost = 0
    self.waiting_time = 0
    self.waiting_cost = 0
    #This variable is for the patients who have not been diagnosed and they are supposed to be referred.
    self.ND = []
    self.diagnosed_status = "ND"    #D: Diagnosed   ND: Not Diagnosed
    self.living_status = "A"   #A: Alive    D:Dead
    self.dead_patient = 0
    self.diagnosed_location = None
```

Figure 30: Python code for patient agent class.

Figure 30 shows how we define the PatientAgent class. Parameters for the PatientAgent class include entry time, diagnosed_status, and others.

```
#Initilizing Agents
for unique_id in range(1, self.num_p + 1):
  patient = Patient(unique_id, self)
  self.schedule.add(patient)
  patient.entry_time = random.randint(0, self.max_steps)
  # Place the agent at a random continuous location
  x = random.uniform(0, self.space.x_max)
  y = random.uniform(0, self.space.y_max)
  # place_agent method adds the coordinate to the agent automatically
  self.space.place_agent(patient, (x, y))
```

Figure 31: Code for initializing patient agent class

Figure 31 shows the code for initializing patient agent including the location of patients which has been generated randomly in a continuous space. Initializing hospitals are the same as the patient's agent as we have explained the hospital initialization in chapter 3.

Here is the map of hospitals and patients' position:

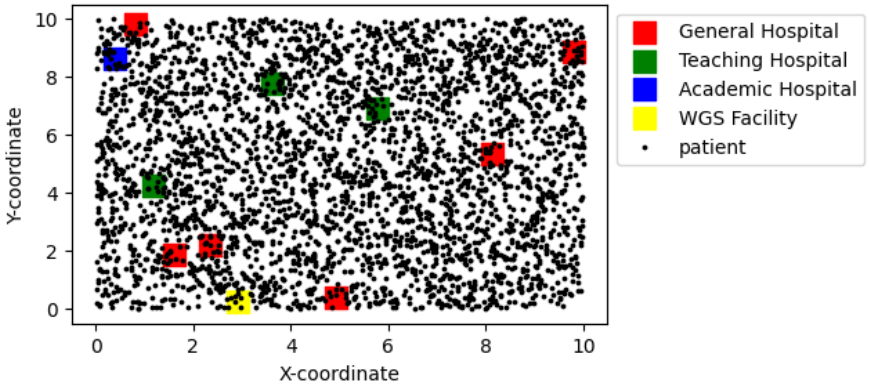

Figure 32: Map of hospital & patients' positions.

Figure 32 shows the randomly generated locations of various hospitals. The red ball indicates general hospitals, marked at six different locations. The green ball represents teaching hospitals, with three distinct locations. The blue and yellow balls indicate the locations of an academic hospital and a WGS facility, respectively, each at a single location. Black points represent patients who randomly visit hospitals.

4.2.2 Comparing diagnosed patients for both scenarios

Here is the graph for comparing number of total diagnosed patients for both scenarios:

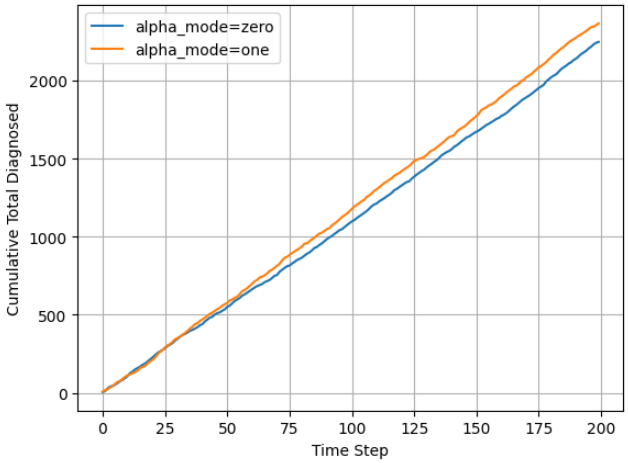

Figure 33: Comparative analysis diagnosed patients for two scenarios

Figure 33 indicates that the base scenario (alpha_mode = one) results in a higher cumulative number of diagnosed patients compared to the scenario 2 (alpha_mode = zero). The time step unit in this simulation is considered to be a day. This means that 200-time steps are equal to 200 days.

4.2.3 Comparing dead patients for both scenarios

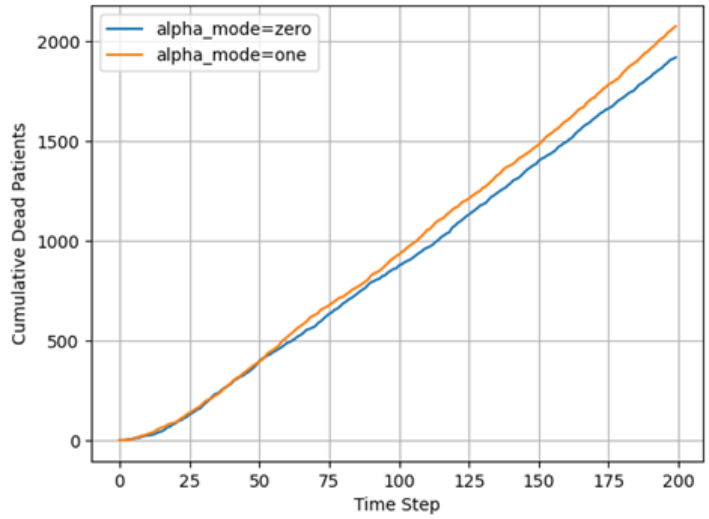

Figure 34: Comparative analysis of dead patients for two scenarios.

Figure 34 displays the cumulative number of dead patients over time for two different scenarios: labeled "alpha_mode = zero"(scenario 2) and "alpha_mode = one" (base scenario). The graph

illustrates that the choice of alpha mode influences the cumulative number of dead patients over time. Specifically, the base scenario results in a higher number of dead patients compared to scenario 2. The number of dead patient for the base scenario is 2041 and the number of dead patients for the scenario 2 is 1985.

4.2.4 Comparing total cost for both scenarios

Before diving into the graph, the method for calculating the total cost is explained below. The mathematical formulas used to calculate the cost are provided here. First, we calculate the cos per patient as below:

$$P_j = Cost\ of\ each\ Patient \qquad j \in J$$

$$P_j = \ cost\ of\ test\ performed\ on\ patient\ j + cost\ of\ waiting\ for\ patient + \ cost\ of\ death\ if\ the\ patient\ died\ in\ the\ middle\ of\ diagnostic\ process.$$

Example:

Patient 1 visited General Hospital 1 without being diagnosed, referred to teaching hospital 2 and being diagnosed at teaching hospital 2.

Cost of tests performed on patient number 1 = 334.14 + 536.02 = 870.16

Cost of waiting for patient 1 = 100 * 30 = 3000

Cost of Death = 0

Cost of each patient (P_j) = 870.16 + 3000 + 0 = 3870.16

Then we calculate cost for hospital based on the cost of each patient:

$$C_i = Cost\ of\ each\ Hospital \qquad i \in I$$

$$P_j = Cost\ of\ Patient\ j \qquad j \in J_i$$

$$J_i = Set\ of\ patients\ visiting\ hospital\ i\ for\ the\ first\ time \quad i \in I$$

$$C_i = \sum_{j \in J_i} P_j$$

Here is the graph for comparing the total cost for both scenarios:

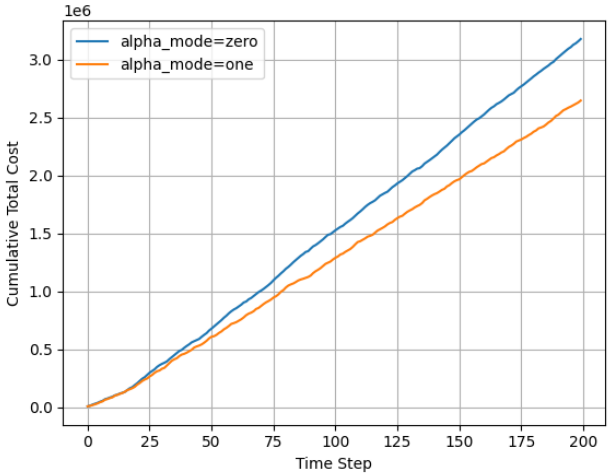

Figure 35: Comparison of cumulative total cost for two scenarios

Figure 35 presented a comparison of the cumulative total cost over time for two different alpha modes, labeled alpha_mode = zero (scenario 2) and alpha_mode = one (base scenario). The graph shows that cumulative total cost of scenario 2 is higher than the base scenario.

4.2.5 Comparing ratios for both scenarios

Here is the graph for comparing base scenario and scenario 2:

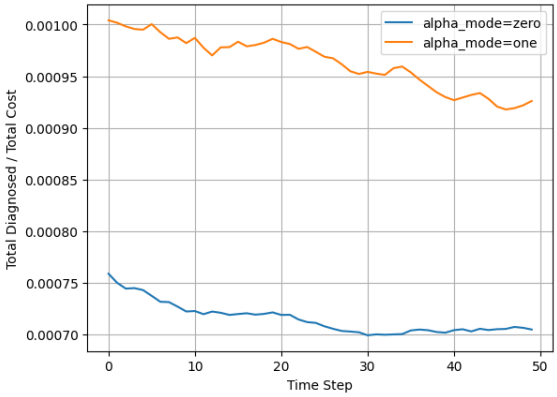

Figure 36: Comparative analysis of ratios of two scenarios.

Figure 36 indicates that the base scenario maintains a higher ratio of diagnosed patients to total cost, implying greater efficiency in terms of cost-effectiveness.

4.3 Comparing scenarios with Cost per Death & Waiting Time

In this section, we compare the results for the base scenario and scenario 2, considering the death rate and associated death cost for patients. As explained in Section 4.1.2.2, the daily death rate (DDR) is 0.0573. Additionally, we have factored in a death cost of 100000 per patient and cost of waiting time 100 per day, and WGS capacity is equal to 4 per day.

4.3.1 Comparing diagnosed patients for both scenarios under waiting time cost and death cost

(a) (b)

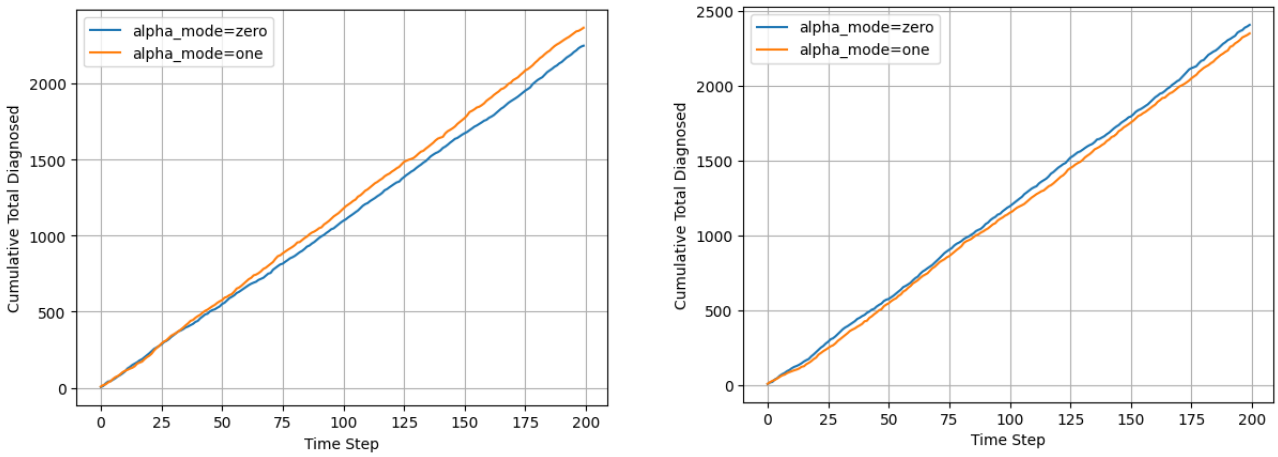

Figure 37: Comparative analysis of diagnosed patients – without death and waiting cost and (a) & with death and waiting cost (b)

Figure 37 presented a comparison of the cumulative total diagnosed patients over time for two different alpha modes, labeled alpha_mode = zero (scenario 2) and alpha_mode = one (base scenario). The graph illustrates that when we consider a cost for death and cost for waiting time, the scenario 2 has a greater number of diagnosed patients.

4.3.3 Comparing total cost for both scenarios under waiting time cost and death cost

(a) (b)

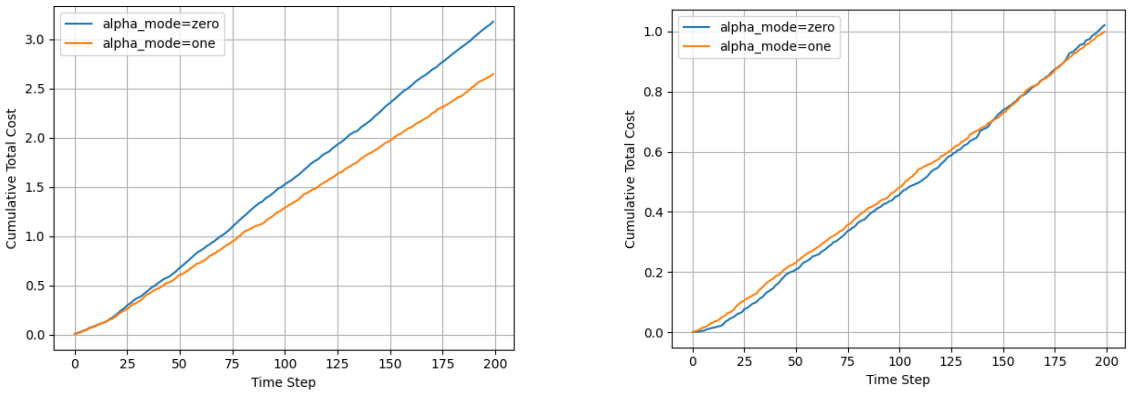

Figure 38: Comparative analysis of cost for two scenarios – without death and waiting cost and (a) & with death and waiting cost (b)

Figure 38 displays the cumulative total cost over time for two different scenarios labeled "alpha_mode = zero"(scenario 2) and "alpha_mode = one" (base scenario). The graph provides a clear comparison of the cumulative total cost for two different alpha modes. Despite the different settings (alpha_mode=zero and alpha_mode=one), the results are very similar, indicating that this particular parameter does not have a substantial effect on the cumulative total cost in the simulation.

4.3.4 Comparing ratios for both scenarios

(a) (b)

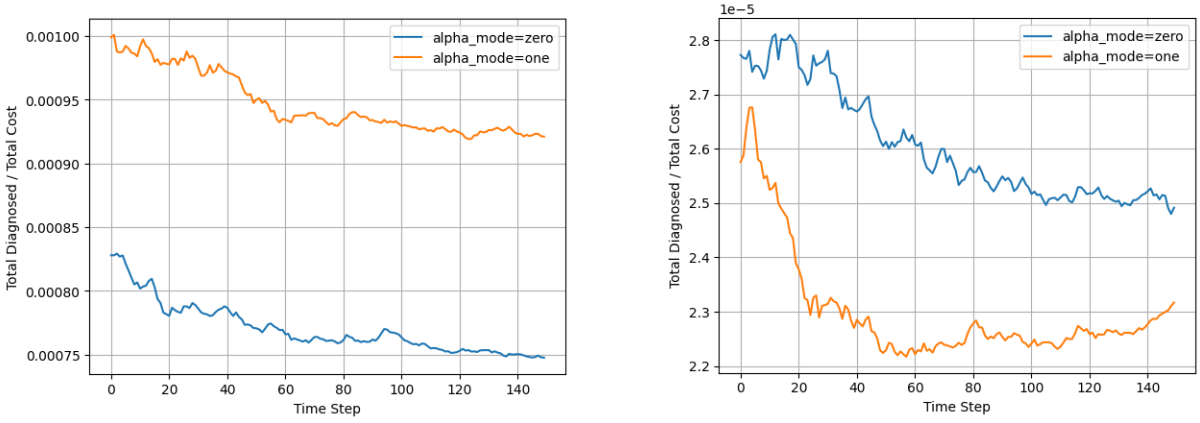

Figure 39: Comparative analysis of ratios – without death and waiting cost (a) and with death and waiting cost (b)

Figure 39 presents a comparison of the ratio of the total diagnosed to the total cost over time for two different alpha modes, labeled alpha_mode = zero (scenario 2) and alpha_mode = one (base scenario). Figure 38 (a) presents the ratio without considering death and waiting cost which in this situation, the base scenario is better. Figure 38 (b) shows the ratio with considering death and waiting cost per patient and in this situation the scenario 2 is better which is sending patient to academic hospital directly from general hospital. It means, when death and waiting have costs, it is better to send patients to academic hospital directly from general hospital.

4.4 Sensitivity Analysis on WGS Capacity

For both the base scenario and scenario 2, we performed a sensitivity analysis on the WGS capacity, varying it from 4 to 1000. Here is the result:

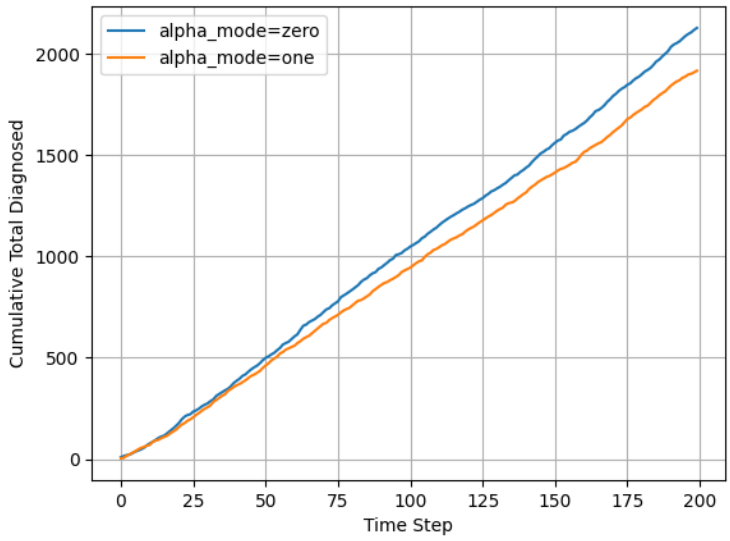

Figure 40: Comparative analysis of diagnosed patients for two scenarios.

Figure 40 presents the cumulative total diagnosed patients over time for two different alpha modes, labeled alpha_mode = zero(scenario 2) and alpha_mode = one (base scenario). When we changed the WGS capacity to 1000, scenario 2 has a greater number of diagnosed patients in comparison with the base scenario. The number of diagnosed patients for the base scenario us 1915 and for the

scenario 2 is 2196 which both are more than the diagnosed patients when the WGS Capacity was 4.

4.5 Conclusion

After a series of analysis and implementing death rate with cost and also considering a cost for waiting time which make the model more realistic, we concluded that scenario 2 which is direct referral from general hospital to academic hospital works better, when we consider a cost for death and waiting time.

CHAPTER V: AGENT BASED MODEL AND APPLICATION OF GAME THEORY

5.1 Scenarios with different alpha values

5.1.1 Comparing diagnosed patients with different alpha values

We have considered different alpha values and run the model to see which alpha would be better. Alpha is a parameter which determines the proportion of patients for sending to teaching or academic hospitals. Alpha is not only zero and one in this analysis, we want to see the other value as well between zero and one.

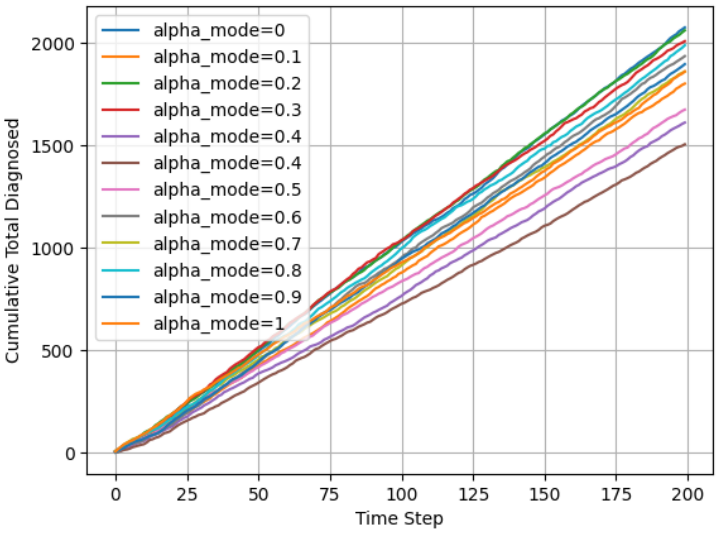

Figure 41: Comparative analysis of diagnosed patients for different alpha.

Figure 41 depicts the cumulative number of diagnosed patients over a time span of 200 steps, with each step representing one day. The model runs simulations for various alpha modes ranging from 0 to 1. This graph effectively illustrates how varying the alpha mode impacts the cumulative number of diagnosed patients over time in the simulation. By comparing these trends, researchers can analyze the influence of different model parameters and configurations on patient diagnosis rates. The distinct colors and labels help in easily distinguishing and comparing the results across different alpha modes.

5.1.2 Comparing ratios with different alpha values

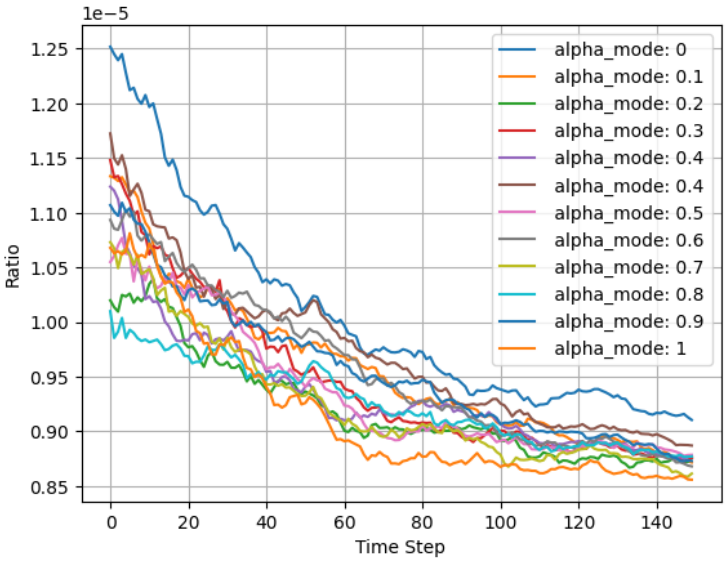

Figure 42: Comparative analysis of ratios for different alpha values.

Figure 42 illustrates how the ratio of diagnosed patients to the total cost evolves over a series of time steps under different alpha mode configurations. This metric provides insight into the efficiency and cost-effectiveness of diagnosing patients over time. his graph effectively demonstrates how the cost-effectiveness of diagnosing patients changes over time under different alpha mode settings. By comparing these trends, researchers can identify which configurations offer the best balance between diagnosing a high number of patients and maintaining manageable costs.

5.2 Cost-Based Alpha Adjustment for General Hospitals Based on Evolutionary Game Theory

In this model, we have generated a random alpha for each General Hospital. Alpha shows the percentage of the individuals who are sent to teaching hospital after visiting a general hospital. For example, we have six General Hospitals, each with a different alpha. In the model class, when we initialize agents, we define alpha for the General Hospital class. When patients proceed to each hospital, we calculate the cost incurred at that hospital. Additionally, if a patient dies, we add the

cost of death, which is $100,000 for that patient. The same applies to the cost associated with waiting time. This calculation continues until the patient is diagnosed. We then calculate the minimum cost, and for each General hospital with the minimum cost, we obtain the alpha of that hospital which has the minimum cost, and we call that alpha a_{min}. The new alpha is calculated using the formula mentioned below:

$I: Sets\ of$ General Hospitals $\quad i = \{1, 2, 3, 4, 5, 6\}$

$C_i: Cost\ per\ General\ Hospitals \quad i \in I$

$CP_i: Cost\ per\ patient \quad i \in I \quad CP_i = \frac{C_i}{p}$

$$\lambda_i = \frac{1}{e^{-(C_i/p\ -\ minC_i)}}$$

$$\alpha^* = \lambda_i a_i \ + \ (1 - \lambda_i) a_{min}$$

We then use this new alpha for other General Hospital agents. We also ran the model with alpha values of 1 and 0, as in Chapter III. Additionally, we proposed a new alpha adjustment to determine which approach works better.

```
#Initialize an empty list to store costs and associated unique_ids
self.hospitals_cost()
costs = []
#Iterate through all agents in the schedule
for agent in self.schedule.agents :
  if agent.agent_type == "GeneralHospital":
    # Append cost and unique_id if the cost is greater than zero
    if agent.initial_patients_cost > 0 :
     costs.append((agent.initial_patients_cost / agent.num_initial_p, agent.unique_id, agent.Alpha))
print('patient_cost:', costs)

# Find the minimum cost among the valid cost
if costs:
  min_cost, min_cost_unique_id, min_alpha = min(costs, key=lambda x: x[0])
  print('Minimum patient cost:', min_cost, 'Unique ID of minimum cost patient:', min_cost_unique_id, 'MIN ALPHA:', min_alpha)
else:
  print('No valid patient costs found.')

#for loop for agents which are general hospital, calculate lambda value and then calculate alpha
for agent in self.schedule.agents:
  if agent.agent_type == "GeneralHospital" and agent.initial_patients_cost > 0:
    lambda_value =  1 / (1 + np.exp(-(agent.initial_patients_cost / agent.num_initial_p  -  min_cost)))
    agent.Alpha = lambda_value * agent.Alpha + (1 - lambda_value) * min_alpha
```

Figure 43: Code for the game theory and cost adjustment

Figure 43 shows how we have defined and coded the mathematical formula explained in section 5.2.

5.2.1 Model with Game Theory logic

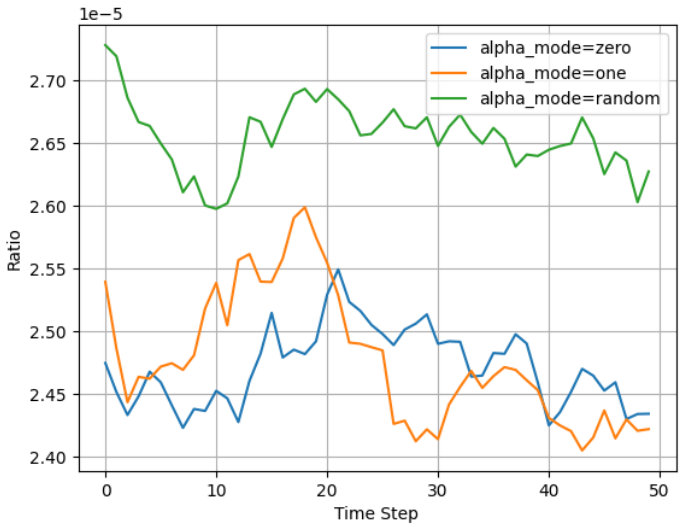

Figure 44: Analysis of total diagnosed to total cost for alpha equal to one, zero and random alpha.

Figure 44 shows the cumulative total diagnosed to cumulative total cost for alpha equal to zero, one and random alpha. Alpha equal to zero is scenario 2 and alpha equal to one is the base scenario.

The green line shows the ratio for random alpha which is better than the base scenario and scenario 2. This model has been run for 100 steps and we got the result excluding the first 50 steps.

5.2.2 Alpha Analysis

As we explained before, alpha is a parameter that determines the proportion of patients sent to teaching or academic hospitals. With this game theory model, we can find the alpha that results in more diagnosed patients at a lower cost. We have also considered WGS capacity, as well as the costs associated with death and waiting time. Here is a graph that shows how alpha converges to a specific value.

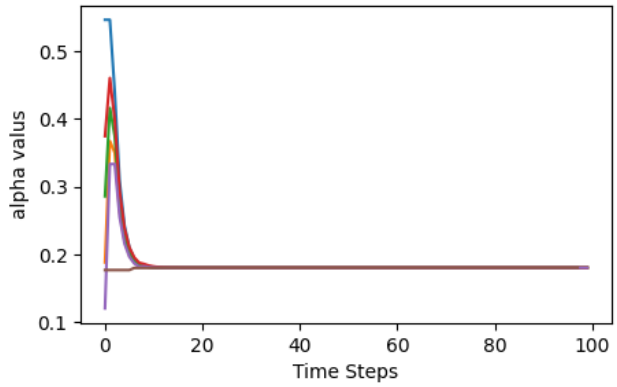

Figure 45:Alpha convergence

Figure 45 shows that alpha converges to 0.2, which means the proportion for sending patients to teaching hospital should be 20% and 80% of patients should be sent to academic hospital.

5.3 Conclusion

In chapter V, we explored various alpha values to determine the optimal configuration for distributing patients between teaching and academic hospitals. Unlike a binary approach, we considered a range of alpha values between 0 and 1 to understand their impacts more comprehensively.

CHAPTER VI: CONCLUSION

This thesis has examined the implementation and impact of WGS in lung cancer diagnosis, utilizing innovative methodologies such as ABS and game theory to optimize healthcare strategies. WGS offers a comprehensive mapping of genetic information, identifying genetic variations linked to diseases like cancer and enabling personalized treatment approaches. Despite its limited use in the USA due to insurance complexities, WGS remains a valuable diagnostic tool in advanced healthcare settings.

The study aimed to develop diagnostic policies that minimize costs and maximize the number of diagnosed patients by integrating ABS and game theory. These methodologies provide a robust framework for analyzing the intricate interactions between hospitals, conceptualizing each as an autonomous agent with distinct behaviors and interaction patterns. Through this analysis, it was found that direct referral from general hospitals to academic hospitals results in a higher number of diagnosed patients when the WGS capacity is 32 per day. By incorporating costs associated with death and waiting time, the model achieved greater realism, demonstrating the effectiveness of this referral strategy under these conditions.

Additionally, the exploration of various alpha values provided insights into the optimal configuration for distributing patients between teaching and academic hospitals. By considering a range of alpha values between 0 and 1, the study identified the most efficient patient distribution strategy, balancing diagnosis rates and associated costs.

In conclusion, this thesis demonstrated that integrating WGS with ABS and game theory can significantly enhance lung cancer diagnosis and treatment. By developing effective diagnostic policies that consider cost factors and patient distribution strategies, the study contributes valuable

insights for optimizing healthcare systems. The findings underscore the potential of WGS as a powerful diagnostic tool and the utility of advanced simulation techniques in improving healthcare outcomes.

REFERENCES

[1] M. Soltanolkottabi, H. A. Khorshidi, and M. J. Ijzerman, “Modeling of Whole Genomic Sequencing Implementation using System Dynamics and Game Theory,” 2022.

[2] M. van de Ven, M. IJzerman, V. Retèl, W. van Harten, and H. Koffijberg, “Developing a dynamic simulation model to support the nationwide implementation of whole genome sequencing in lung cancer,” *BMC Med Res Methodol*, vol. 22, no. 1, p. 83, Dec. 2022, doi: 10.1186/s12874-022-01571-3.

[3] “What is Whole Genome Sequencing?” Accessed: Oct. 06, 2023. [Online]. Available: https://www.illumina.com/techniques/sequencing/dna-sequencing/whole-genome-sequencing.html

[4] “What is a genome?” Accessed: Oct. 06, 2023. [Online]. Available: https://www.yourgenome.org/facts/what-is-a-genome/

[5] “What is WGS?” Accessed: Oct. 06, 2023. [Online]. Available: https://sequencing.com/blog/post/what-is-whole-genome-sequencing-wgs

[6] Mohapatra, “Application of Biomarkers in Cancer Diagnosis and Treatment”.

[7] “Lungevity.” Accessed: Aug. 07, 2024. [Online]. Available: https://www.lvng.com/resources.html?source=tag_c_c_1035&umedium=cpc&uadpub=bing&ucampaign=2022lvngwith_linkedintest_dtc&ucreative=linkedin_resources_ph&uplace=lungevity&outcome=udtc&cmpid=1&&msclkid=254d6b84746615d8b6f10d315e48bb40&gclid=254d6b84746615d8b6f10d315e48bb40&gclsrc=3p.ds

[8] Yinfa Ma & Sanjeewa Gamagedara, “Biomarker analysis for oncology”.

[9] A. J. Aarti Desai, “Next-Generation Sequencing for Cancer Biomarker Discovery”.

[10] K. R. Loeb and L. A. Loeb, “Significance of multiple mutations in cancer,” 2000.

[11] K. Schwarze *et al.*, “The complete costs of genome sequencing: a microcosting study in cancer and rare diseases from a single center in the United Kingdom,” 2020, doi: 10.1038/s41436.

[12] “What is lung cancer?” Accessed: Oct. 06, 2023. [Online]. Available: https://www.cdc.gov/cancer/lung/basic_info/what-is-lung-cancer.htm

[13] K. Schwarze, J. Buchanan, J. C. Taylor, and S. Wordsworth, “Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature,” Oct. 01, 2018, *Nature Publishing Group*. doi: 10.1038/gim.2017.247.

[14] L. G. Gordon *et al.*, “Estimating the costs of genomic sequencing in cancer control,” *BMC Health Serv Res*, vol. 20, no. 1, p. 492, Dec. 2020, doi: 10.1186/s12913-020-05318-y.

[15] “WGS in Yale.” Accessed: Oct. 25, 2023. [Online]. Available: https://www.yalemedicine.org/conditions/whole-genome-sequencing

[16] “Introductory tutorial: Agent-based modeling and simulation,” Savannah, GA, USA: IEEE.

[17] C. M. MacAl and M. J. North, “Tutorial on agent-based modelling and simulation,” *Journal of Simulation*, vol. 4, no. 3, pp. 151–162, 2010, doi: 10.1057/jos.2010.3.

[18] M. R. Friesen and R. D. McLeod, “A survey of agent-based modeling of hospital environments,” *IEEE Access*, vol. 2, pp. 227–233, 2014, doi: 10.1109/ACCESS.2014.2313957.

[19] Sean Barnes, Bruce Golden, and Stuart Price, “Applications of Agent-Based Modeling and Simulation to Healthcare Operations Management,” Springer, New York, NY.

[20] “An Introduction to Game Theory”.

[21] R. R. Hill, L. E. Champagne, and J. C. Price, “Using Agent-based Simulation and Game Theory to Examine the WWII Bay of Biscay U-boat Campaign,” 2004.

[22] M. Soltanolkottabi, “Modeling social response to disease spread using spatial game theory,” AN ABSTRACT OF A DISSERTATION, 2008.

[23] “Fundamentals of Evolutionary Game Theory and its Applications.” [Online]. Available: http://www.springer.com/series/11930

[24] K. Chatterjee, R. Ibsen-Jensen, I. Jecker, and J. Svoboda, “Complexity of Spatial Games,” in *Leibniz International Proceedings in Informatics, LIPIcs*, Schloss Dagstuhl- Leibniz-Zentrum fur Informatik GmbH, Dagstuhl Publishing, Dec. 2022. doi: 10.4230/LIPIcs.FSTTCS.2022.11.

[25] M. A. NOWAK, *Evolutionary Dynamics*. Harvard University Press, 2006. doi: 10.2307/j.ctvjghw98.

[26] “Everyday Health.” Accessed: Jul. 03, 2024. [Online]. Available: https://www.everydayhealth.com/lung-cancer/living-with-stage-4-lung-cancer-survival-rates-treatments-emotional-support/

APPENDICES

Google Colab Code

You can access the Google Colab code for this project at the following link:

Chapter III Model:

https://github.com/Hastighn/Lung-cancer-treatment--agent-based-model

Chapter IV Model:

https://github.com/Hastighn/lung-cancer-treatment-agent-based-modeling/tree/main