FNAS Journal of Mathematical Modeling and Numerical Simulation
Print ISSN: 3027-1282
www.fnasjournals.com
Volume 2; Issue 4; December 2025; Page No. 31-44.
DOI: https://doi.org/10.63561/jmns.v2i4.1121

Stochastic Game Models Based on Markov Decision Processes for Multi-Agent Decision Making in Clinical Healthcare Systems

***[1]Udok, U.V., [2]Victor-Edema, U. A., & [3]Ijomah, M.A.**
[1]University of Uyo Teaching Hospital, Uyo. Akwa Ibom State
[2]Department of Mathematics, Ignatius Ajuru University of Education, Port Harcourt
[3]Department of Statistics University of Port Harcourt. Rivers State

***Corresponding author email:** udemevin@gmail.com

Abstract
The application of stochastic game models on Markov Decision Processes for multi-agent decision making in clinical healthcare settings has spiked some interests in recent times. Despite advancements in single-agent models, there remains a notable knowledge gap in incorporating multi-agent strategic interactions within stochastic frameworks that adequately address uncertainty in simulation-based environment. This study addressed a preliminary stochastic analysis of patient progression through distinct health states, influenced by healthcare interventions and a simulated patient summary table. The "Recovered" state was identified as an absorbing state, consistent with observed final health outcomes where all patients eventually recovered. The analysis further considered two principal agents: the patient, characterized by initial severity (Mild, Moderate, Severe) and risk (Low, Medium, High) which directly influenced their initial state and potential health progression. The study adopted a simulated-based analysis framework implemented in Python. The healthcare system/decision-makers, whose "Treat" or "Wait" interventions were hypothesized to impact the stochastic transitions. Key results demonstrated that a consistently applied "Treat" policy effectively guided patients through defined health states towards a recovered absorbing state, evidenced by high transition probabilities towards improved conditions. The computed value function, the expected reward for each state, derived via the Bellman equation with a discount factor of 0.95, revealed that managing patients from a "Critical" state, a moderate average payoff of 6.8, improving recovery odds by 15-20%. The framework applied Stochastic game theory based on Markov Decision Processes, the framework posited that "Treat" decisions would accelerate positive transitions and minimized negative ones, offering a higher patient payoff. Two (4 ×4) matrices were created from simulated-based data for transition probability on "treat" and "wait" actions. Markov model was constructed, capturing transitions between health states: Critical, Serious, Stable, and Recovered. Transition probability matrices revealed that all states eventually absorbed into Recovery with probability 1.0. The expected time to recovery from Critical was 3.25 compared to 6.0 from Serious and 7.75 from Stable. Using a reward structure penalizing critical states (0) and rewarding recovery (+8.5), the expected cumulative reward from each initial state was computed as: Critical = 3.25, Serious = 6.0, Stable = 7.75. However, the study recommended that healthcare agents and systems should improve clinical decision-making under uncertainty by applying Markov Decision Processes to minimize patient times spent in critical or serious health states, delays and costs of care in order to ensure evidenced-based support and overall system performance.

Keywords: Stochastic game theory, Markov decision processes, Multi-agent, Decision-making, Healthcare

Introduction
A Stochastic Game Theory framework, also known as a Markov Game, provides a sophisticated approach to modeling complex, dynamic decision-making scenarios in clinical healthcare involving multiple interacting agents where the

consequences of actions unfold probabilistically over time. The framework allows for the analysis of strategic interactions where each agent's decisions not only yield immediate rewards or costs but also influence the probability of transitioning to future states of the system, typically representing the evolving health status of patients or the operational conditions of a healthcare system. Stochastic game theory provides a formal structure for analyzing such scenarios by modeling agents' interactions as probabilistic state transitions influenced by their collective strategies (Basar & Olsder, 1999). This makes it particularly suitable for modeling real-world healthcare problems such as patient flow optimization, treatment planning, and resource allocation. Game theory has been applied in various fields such as social science, economics, medical science, political science, and management science. It started to gain popularity in medical decision-making, doctor–patient interactions, organ transplant management, resource planning, and training, (McFadden, 2012). Zhu et al. (2016) developed a framework based on evolutionary game theory that combines group phenotypic composition with ecological interactions. Their framework specifically maps Quantitative Trait Loci (QTLs) for population demographics and evolution. Archetti (2013) employed evolutionary game theory to study the collective interaction between cancer cells, analyzing the dynamics of these cells' growth factors and treatment effectiveness in reducing the cell population. Game theory has shown successful application in medical resource management and training, yielding favourable outcomes. McFadden and Tsai et al. (2012) applied game theory in complex operating room system management, resulting in positive effects on the environment and benefiting all stakeholders. Blake and Carroll (2016) proposed using game theory in medical training and practice to encourage better recognition of competing priorities and adjustment of approaches when one's preferred outcome is unlikely

Clinical healthcare settings are inherently complex systems involving a multitude of interacting agents, including patients, physicians, nurses, specialists, hospitals, payers, and policymakers. These agents make sequential decisions under conditions of significant uncertainty, where the outcomes of their choices, such as treatment effectiveness, disease progression, and resource utilization, are often probabilistic (Folland et al., 2017). Therefore, this study aims to fill this gap by developing a stochastic game theory framework for multi-agent decision-making in clinical healthcare settings, providing MDP theoretical underpinnings and practical insights for enhancing health outcomes.

This chapter reviews the related literature on stochastic game models on Markov Decision Processes for multi-agent decision making in clinical healthcare settings. The theoretical and conceptual framework serves as the foundation upon which this study is grounded. For a comprehensive understanding of the implementation of Multi-Agent Systems (MAS) within healthcare, particularly under conditions of uncertainty, it is necessary to draw from multiple theories and conceptual underpinnings. These include Decision Theory, Game Theory (particularly Stochastic Game Theory), Complex Adaptive Systems Theory, Agent-Based Modeling, and a derived conceptual framework that guides the empirical implementation of the study. Game theory studies interactions among rational decision-makers. It is highly applicable in MAS where multiple agents, each with their own objectives, must interact, cooperate, or compete. In healthcare, game theory has been applied to model negotiations between stakeholders, resource allocation, and treatment planning (Osborne & Rubinstein, 1994).

Myerson (2004), extended classical models with incomplete information to Bayesian games, providing a foundation for decision-making under uncertainty. In healthcare, these models are used to optimize the interactions between healthcare providers and patients (or among providers) in cost-sharing, prescription behavior, or compliance monitoring (Mendonça et al., 2020). In more advanced scenarios, stochastic games represent an extension of repeated games with probabilistic transitions that capture the temporal and uncertain nature of clinical interactions (Maski & Tirole 2001). These models allow decision-makers to plan over time while considering future consequences, making them suitable for chronic disease management and hospital admission strategies, a study by Adida et al. (2018) employed a stochastic game model to analyze patient adherence and physician effort in managing coronary heart disease, considering behavioral factors. These models highlight how patient choices and provider interventions dynamically influence the patient's health trajectory.

In Public Health and Epidemiology, stochastic game theory, including its mean-field game variations, has been applied to model the spread of infectious diseases and the strategic responses of individuals to public health interventions like vaccination (Reluga, 2011) Healthcare decision-making, an important process in which the best action to achieve the desired goals is chosen, largely determines the quality of care, patient safety, and the possibility of future

Cite this article as:
Udok, U.V., Victor-Edema, U. A., & Ijomah, M.A. (2025). Stochastic game models based on markov decision processes for multi-agent decision making in clinical healthcare systems. *FNAS Journal of Mathematical Modeling and Numerical Simulation*, *2*(4), 31-44. https://doi.org/10.63561/jmns.v2i4.1121

complications, (Stubbings et al.,2012). As an essential part of the professional duties of the medical personnel, clinical decision-making consists of analysis of information, making decisions, and taking action based on those decisions to accomplish the desired objective, (Wu et al., 2016). In other words, game theory deals with mathematical models of cooperation and conflicts between rational decision-makers. Game theory can be defined as the study of decision-making in which the players must make strategies affecting the interests of other players. Multi-actor decision-making is complex, given the involvement of multiple actors whose behaviour and interactions steer the process of decision-making, (De Bruijn & Heuvelhof, 2018). Decision-making situations can be viewed as games in examining strategic behaviours and interactions (Scharpf, 1997). The versatility of game theory has led to diverse healthcare applications including: Epidemiology: Models have been used to analyze vaccination behavior, quarantine strategies, and treatment adoption in epidemics. For instance, Bauch and Earn (2004) used game theory to show how individual choices in vaccination uptake can lead to suboptimal population-level immunity. Game theory helps insurers design contracts that minimize adverse selection and moral hazard, aligning incentives among patients, providers, and payers. Telemedicine and E-Health Services:

Game theory provides a powerful mathematical tool to model and analyze strategic interactions between rational decision-makers. In healthcare, it has been applied to problems such as organ allocation, insurance design and resource competition However, classical game theory assumes static and fully observable environments, which do not reflect the real-world dynamics of patient care. Conversely, stochastic models, including Markov Decision Processes (MDPs), allow for the modeling of uncertainty in disease evolution but typically involve a single decision-maker and fail to capture the strategic interplay between multiple agents. The stochastic game theory framework which merges game theory with stochastic processes, offers a promising solution by enabling the analysis of multi-agent strategic decision-making in uncertain, dynamic environments. Yet, its application in clinical decision-making remains minimal (Schelling, 2010). More recent work has begun to explicitly model the sequential nature of healthcare decisions using stochastic game paradigms. For instance, models have been developed to understand sequential treatment choices for chronic diseases, where the "game" unfolds over multiple patient visits (Long et al., 2017). A core strength of stochastic game theory lies in its capacity to delineate and analyze the interplay between multiple decision-makers. Healthcare agents, including patients, doctors, nurses, and administrators, often possess distinct utility functions and information sets, leading to a blend of cooperative and competitive dynamics.

Cooperative game theory illuminates scenarios where agents can form coalitions and collaborate towards a common goal, such as a surgical team working synergistically to optimize a patient's surgical outcome (Liu et al., 2011). In such instances, concepts like the core or Shapley value can be employed to fairly distribute the benefits of cooperation or assign responsibility. For example, a study might model the cooperative decision-making of a multidisciplinary tumor board, where oncologists, radiologists, and pathologists collaboratively determine the optimal treatment plan for a cancer patient (Rubin et al., 2016). Conversely, non-cooperative game theory is essential for understanding situations where agents act in their own self-interest, potentially leading to competition for scarce resources or differing preferences. This can manifest as competition among hospitals for patient market share, departments vying for budget allocations, or even a patient's non-adherence to treatment due to perceived inconvenience, conflicting with a doctor's recommendation (Mohr & Bittar, 2014). The identification of Nash Equilibria in these non-cooperative settings is critical for predicting stable outcomes where no agent has an incentive to unilaterally deviate from their chosen strategy. Furthermore, the concept of Markov Perfect Equilibrium (MPE) is particularly pertinent in stochastic healthcare games, as it assumes strategies depend only on the current state of the system, reflecting the real-time, adaptive nature of clinical decisions without relying on full historical information (Maskin & Tirole, 2001).

Acuna et al. (2021) developed a stochastic game theory framework for multi-agent decision-making in clinical healthcare settings offers a powerful and comprehensive approach to understanding and optimizing complex interactions. By integrating the dynamic nature of patient health states, the interplay of cooperative and competitive agent behaviors, and the pervasive presence of uncertainty through MDPs. The reliance on simulation-based analysis further ensures that the insights generated are robust and relevant to the practical constraints of real-world clinical environments. The collaborative approach of shared decision-making can lead to more patient-centered care and improve treatment adherence and outcomes (Vahdati et al., 2024). Patients may value healthcare providers' expertise in guiding treatment decisions with specialized knowledge that is essential in building trust and confidence in them.

Cite this article as:
Udok, U.V., Victor-Edema, U. A., & Ijomah, M.A. (2025). Stochastic game models based on markov decision processes for multi-agent decision making in clinical healthcare systems. *FNAS Journal of Mathematical Modeling and Numerical Simulation, 2*(4), 31-44. https://doi.org/10.63561/jmns.v2i4.1121

Moreover, this participation causes improved control of diabetes, better physical functioning in rheumatic diseases, enhanced patients' compliance with secondary preventive actions and improvement in health of patients with myocardial infarction, (Arnetz et al., 2007). Predictive models for sepsis, readmissions, and treatment response now aid clinicians in real-time, often embedded within clinical decision support tools. Reinforcement learning, a branch of AI related to stochastic game theory, is increasingly used to personalize care pathways, particularly in intensive care and oncology (Komorowski et al., 2018). Despite promising results, concerns regarding data privacy, model interpretability, and ethical use remain critical barriers to full adoption. Nonetheless, these models represent the frontier of modern decision-making in healthcare. With the increasing complexity of healthcare delivery, involving multiple stakeholders such as physicians, nurses, patients, caregivers, and administrative personnel has made decision-making evolved into a multi-agent process. Multi-Agent Systems (MAS) involve autonomous entities (agents) that interact to achieve individual or collective goals, and are increasingly used to model dynamic, decentralized healthcare environments (Jennings, et al, 1998). It refers to a collection of autonomous, intelligent agents that interact with one another within an environment to achieve individual or collective goals. In healthcare, these agents may represent software systems, robots, clinical decision-support tools, or human professionals like doctors and nurses.

A model developed by Klein (1993) demonstrates how agents can be programmed to coordinate test results, drug administration, and ventilator adjustments based on continuously updated patient data. The MAS system reacts dynamically to patient deterioration and alerts human caregivers, improving response time and reducing mortality risk. Markov Perfect Equilibrium (MPE) is a refinement of the Nash equilibrium used in dynamic settings, particularly stochastic games. In MPE, players' strategies depend only on the current state of the game, not the full history of play. In healthcare, MPE can model decision-making over time where the system's state (e.g., patient condition, hospital congestion) evolves. In treating chronic diseases, decisions on medication dosages or interventions depend on the current health state of a patient.

MPE identifies optimal treatment strategies that adapt dynamically over time, maximizing patient outcomes under probabilistic health state transitions (Hauskrecht, 2000). Multi-Agent Systems (MAS) and Game Theory are two prominent frameworks that have seen growing application in healthcare for facilitating decision-making, resource allocation, diagnostics, and treatment planning. One of the most commonly used stochastic models in healthcare is the Markov Decision Process (MDP). MDP models are used to model decision-making in situations where outcomes are partially random and partially under the control of the decision-maker. In the context of healthcare, an MDP provides a structured approach to model healthcare decisions, where the system evolves from one state to another with certain probabilities, and each state is associated with a reward or cost (Puterman, 2005).

An MDP is defined by states, actions, transition probabilities and rewards. Where, states (S) represent the different possible conditions of the system (e.g., different stages of a patient's disease or health condition). Actions (A), represent the decisions or interventions that can be taken (e.g., different treatment options or procedures). Transition Probabilities (P) represent the probabilities of moving from one state to another after taking an action. The immediate reward or cost associated with being in a particular state and taking a specific action (e.g., the cost of a treatment, the improvement in patient health). MDP models are particularly useful in chronic disease management, where a patient's condition evolves over time, and medical interventions can either improve or worsen the patient's health. For example, they can model decisions in managing diabetes, where actions (such as medication or lifestyle changes) lead to transitions between health states (e.g., controlled vs. uncontrolled diabetes) with associated costs and benefits. These models can also incorporate value iteration or policy iteration algorithms to compute the optimal policy and a sequence of actions that maximizes long-term benefits or minimizes long-term costs for the patient, (Filar & Vrieze, 1997). While these frameworks are effective for modeling fully observable systems, many real-world clinical decision-making scenarios involve incomplete information.

Stochastic game theory offers a dynamic modeling framework that can account for the evolving nature of patient care and healthcare systems. By integrating uncertainty and time into the decision-making process, stochastic games provide a more realistic and flexible approach to optimizing decisions over time. Healthcare systems are inherently multi-agent environments where multiple agents (doctors, nurses, patients, insurers, etc.) interact with one another, each with different goals, preferences, and information, (Adida et al., 2018). However, most existing decision-making

Cite this article as:
Udok, U.V., Victor-Edema, U. A., & Ijomah, M.A. (2025). Stochastic game models based on markov decision processes for multi-agent decision making in clinical healthcare systems. *FNAS Journal of Mathematical Modeling and Numerical Simulation, 2*(4), 31-44. https://doi.org/10.63561/jmns.v2i4.1121

frameworks either simplify or ignore these interactions.

Aim and Objectives of the Study:
The aim of this study is to present the application of a Stochastic game theory framework for Multi-agent decision-making in clinical healthcare settings.
The specific objectives of this research are to:
i. develop a Stochastic game-theoretic framework for modeling dynamic, interactive decision-making in clinical healthcare settings.
ii. Incorporate Stochastic game models based on Markov Decision Processes (MDPs) within the developed framework.
iii. evaluate outcomes under realistic clinical constraints, including limited resources, and time-sensitive interventions, using simulation-based analysis.

Materials and Method
This study adopts a simulation-based research design to investigate decision-making in healthcare systems using stochastic game theory and reinforcement learning. The model-developing approach focuses on constructing a framework that represents the dynamic nature of healthcare decisions, incorporating time-dependent variables. Also, healthcare decision making scenario data from specific health conditions are used for the analysis.
It emphasizes the importance of understanding interactions between key healthcare agents, including patients, doctors, nurses, and medical systems, and how these agents can make decisions to achieve both individual and collective goals. Markov decision process, also called stochastic dynamic program or stochastic control problem is a model for sequential decision making when outcomes are uncertain. Model implementation is done in Python (using libraries such as Numpy, TensorFlow, PyTorch, OpenAI Gym, PettingZoo).

Markov Decision Process (MDP) Model
The simulation-based data can be analyzed through the lens of a Markov Decision Process (MDP), which models sequential decision-making for patient health management. The MDP is defined by the tuple (S, A, P, R, γ), where:

- **States** S: The patient health states are $S = \{0,1,2,3\}$, corresponding to Critical (0), Serious (1), Stable (2), and Recovered (3). The Recovered state is absorbing (once reaches, the patient stays there) state.
- **Actions** A: The decisions are $A = \{\text{Treat,Wait}\}$. "Treat" represents active intervention (e.g., medication or surgery), which may improve the state but could have costs or risks. "Wait" represents observation or no intervention, which may allow natural recovery but risks deterioration.
- **Transition Probabilities** P: $P(s' \mid s, a)$ is the probability of transitioning from state s to state s' under action a.
- **Discount Factor** γ: A value like $\gamma = 0.95$ discounts future rewards, emphasizing short-term health improvements

Rewards R: $R(s, a)$ is the immediate reward for taking action a in state s. Reward Function ($R(s, a)$): The immediate reward received for taking action a in state s is defined by a function $R: \mathcal{S} \times \mathcal{A} \to \mathbb{R}$. The values were derived for 'Treat' and hypothesized for 'Wait' as: $R(s,\text{Treat}) = \{1.0,2.0,3.0,0.0\}$ for states $s \in \{\text{Crit, Ser, Stab, Rec}\}$. $R(s,\text{Wait}) = \{-5.0,0.5,4.0,0.0\}$ for states $s \in \{\text{Crit, Ser, Stab, Rec}\}$. A Bellman Equation for Policy Evaluation: The Value Function $V^{\pi}(s)$, representing the expected cumulative reward for a given policy π starting from state s, is defined by the Bellman Equation: $V^{\pi}(s) = R(s, \pi(s)) + \gamma \sum_{s' \in \mathcal{S}} P\,(s'|s, \pi(s)) V^{\pi}(s')$

The MDP component models the patient progression under the assumption of full observability of the patient's true health state. Transition Probability ($P(s'|s, a)$): The probability of transitioning from state s to state s' after taking action a is defined by two 4×4 matrices:

Cite this article as:
Udok, U.V., Victor-Edema, U. A., & Ijomah, M.A. (2025). Stochastic game models based on markov decision processes for multi-agent decision making in clinical healthcare systems. *FNAS Journal of Mathematical Modeling and Numerical Simulation*, *2*(4), 31-44. https://doi.org/10.63561/jmns.v2i4.1121

The optimal Value Function $V^*(s)$, which provides the maximum possible expected cumulative reward starting from state s by choosing the optimal action, is defined by the Bellman Optimality Equation: $V^*(s) = \max_{a\in\mathcal{A}}(R(s,a) + \gamma\sum_{s'\in\mathcal{S}} P(s'|s,a)V^*(s'))$

$$P(s'_t\,|s_t,a) = P(s'_t\,|s_t,a_1,a_2\ldots,a_m) = \frac{N_i}{N} \quad (1)$$

$$R(s,a) = r(s) + c(a) \quad (2)$$

$$r(s) = (10 - severity\ score \times risk\ score) \quad (3)$$

Bellman optimality equation:

$$V_{k+1}(s) = \max_{a\in\{\text{Wait,Treat}\}}[R(s,a) + \gamma\sum_{s'\in S} P(s' \mid s,a)V_k(s')] \quad (4)$$

$$R(s,a) = \sum_{s'} P(s'|s,a)\,\mathrm{r}(s,a,s') \quad (5)$$

$$V(s_t) = \sum_{s_{t+1}\in S} P(s_{t+1} \mid s_t, Treat)[r_{s_t}, s_{t+1} + V(s_{t+1})] \quad (6a)$$

$$V(s_t) = \sum_{s_{t+1}\in S} P(s_{t+1} \mid s_t, Wait)[r_{s_t}, s_{t+1} + V(s_{t+1})] \quad (6b)$$

$$V(s) = R(s,Treat) + \sum_{s'\in S} P(s' \mid s, Treat)V(s') \quad (7a)$$

$$V(s) = R(s,Wait) + \sum_{s'\in S} P(s' \mid s, Wait)V(s') \quad (7b)$$

$$V(s) = R(s,Treat) + P_{s,C}V(C) + P_{s,S}V(S) + P_{s,B}V(B) + P_{s,R}V(R) \quad (8a)$$

$$V(s) = R(s,Wait) + P_{s,C}V(C) + P_{s,S}V(S) + P_{s,B}V(B) + P_{s,R}V(R) \quad (8b)$$

$$V(C) = R(C,Treat) + PC,SV(S) + PC,BV(B) \quad (9a)$$

$$V(C) = R(C,Wait) + PC,SV(S) + PC,BV(B) \quad (9b)$$

$$V(S) = R(S,Treat) + PS,BV(B) + PS,RV(R) \quad (10a)$$

$$V(S) = R(S,Wait) + PS,BV(B) + PS,RV(R) \quad (10b)$$

$$V(B) = R(B,Treat) + PB,BV(B) + PB,RV(R) \quad (11a)$$

$$V(B) = R(B,Wait) + PB,BV(B) + PB,RV(R) \quad (11b)$$

Cite this article as:
Udok, U.V., Victor-Edema, U. A., & Ijomah, M.A. (2025). Stochastic game models based on markov decision processes for multi-agent decision making in clinical healthcare systems. *FNAS Journal of Mathematical Modeling and Numerical Simulation*, *2*(4), 31-44. https://doi.org/10.63561/jmns.v2i4.1121

$$V(R) = R(R, Treat) + PR, RV(R) \quad (12a)$$

$$V(R) = R(R, Wait) + PR, RV(R) \quad (12b)$$

Cite this article as:
Udok, U.V., Victor-Edema, U. A., & Ijomah, M.A. (2025). Stochastic game models based on markov decision processes for multi-agent decision making in clinical healthcare systems. *FNAS Journal of Mathematical Modeling and Numerical Simulation*, *2*(4), 31-44. https://doi.org/10.63561/jmns.v2i4.1121

Results

This chapter presents and interprets the results of the stochastic game model implemented to simulated clinical decision-making involving multiple agents (patients, doctors and healthcare systems). To develop a stochastic game-theoretic framework for modeling dynamic, interactive decision-making in clinical healthcare settings, the key variables are: condition severity: mild, moderate, severe, risk level: low, medium, high and payoff where: Severity Score is grouped into Mild = 1, Moderate = 2, Severe = 3 while Risk Score is categorized into: Low = 1, Medium = 2, High = 3. Each case is assessed across two categorical variables—Severity and Risk—which are quantified into numeric scores (Severity Score and Risk Score respectively). A derived metric, Payoff, serves as an index of overall prognosis or expected outcome. Interpretations are provided to guide potential actions or decisions based on these scores. The table below summarizes the outcomes of the classification model:

Cite this article as:
Udok, U.V., Victor-Edema, U. A., & Ijomah, M.A. (2025). Stochastic game models based on markov decision processes for multi-agent decision making in clinical healthcare systems. *FNAS Journal of Mathematical Modeling and Numerical Simulation*, *2*(4), 31-44. https://doi.org/10.63561/jmns.v2i4.1121

Table 1: Simulated Summary of Each Patient Result

ID	Severity	Risk	Severity score	Risk Score	Payoff	Interpretation
P001	Moderate	Low	2	1	8	Medium condition, low risk – manageable case
P002	Mild	Medium	1	2	8	Mild condition, some risk – likely recoverable
P003	Mild	High	1	3	7	Mild issue, high risk – may need close monitoring
P004	Moderate	Medium	2	2	6	Medium severity and risk – moderately concerning
P005	Mild	Low	1	1	9	Mild condition and low risk – very favorable
P006	Severe	Low	3	1	7	Serious condition but low risk – likely stable
P007	Severe	Low	3	1	7	Serious condition but low risk – likely stable
P008	Severe	High	3	3	1	Very severe, high risk – critical and high alert
P009	Moderate	Medium	2	2	6	Medium severity and risk – moderately concerning
P010	Mild	Low	1	1	9	Mild condition and low risk – very favorable
P011	Moderate	High	2	3	7	Medium severity and high risk – moderately concerning
P012	Mild	High	1	3	7	Mild condition, high risk – needs close monitoring
P013	Moderate	Low	2	1	8	Mild condition, high risk – manageable case
P014	Severe	High	3	3	6	Very severe and high risk – critical and high alert
P015	Severe	Low	3	1	7	Very severe but mild risk – likely stable
P016	Mild	Low	1	1	9	Mild condition and low risk – very favorable
P017	Moderate	Medium	2	2	6	Medium severity and risk – moderately concerning

Cite this article as:
Udok, U.V., Victor-Edema, U. A., & Ijomah, M.A. (2025). Stochastic game models based on markov decision processes for multi-agent decision making in clinical healthcare systems. *FNAS Journal of Mathematical Modeling and Numerical Simulation*, *2*(4), 31-44. https://doi.org/10.63561/jmns.v2i4.1121

P018	Mild	Medium	1	2	8	Mild condition and risk – likely recoverable
P019	Severe	Low	3	1	7	Very severe and low risk – likely stable
P020	Mild	Low	1	1	9	Mild condition and low risk – very favorable

Source: Simulation-Based Data

Table 2: Transition Matrix for "Treat''

From \ To	*Critical*	*Serious*	*Stable*	*Recovered*
Critical	0	17	3	0
Serious	0	0	15	5
Stable	0	0	5	15
Recovered	0	0	0	20

Source: Simulation-Based Data

Table 3 Summary of the Reduced Transition Matrix for ''Treat'' based on Eqn (1)

From \ To	*Critical*	*Serious*	*Stable*	*Recovered*
Critical	0	0.85	0.15	0
Serious	0	0	0.75	0.25
Stable	0	0	0.25	0.75
Recovered	0	0	0	1

Cite this article as:
Udok, U.V., Victor-Edema, U. A., & Ijomah, M.A. (2025). Stochastic game models based on markov decision processes for multi-agent decision making in clinical healthcare systems. *FNAS Journal of Mathematical Modeling and Numerical Simulation*, *2*(4), 31-44. https://doi.org/10.63561/jmns.v2i4.1121

Table 4 Transition Matrix for ''Wait''

From \ To	*Critical*	*Serious*	*Stable*	*Recovered*
Critical	16	2	2	0
Serious	6	10	4	0
Stable	2	4	12	2
Recovered	0	0	0	20

Source: Simulation-Based Data

Table 5 Summary of the Reduced Transition Matrix for ''Wait'' based on Eqn (3.1)

From \ To	*Critical*	*Serious*	*Stable*	*Recovered*
Critical	0.8	0.1	0.1	0
Serious	0.3	0.5	0.2	0
Stable	0.1	0.2	0.6	0.1
Recovered	0	0	0	1

Table 6 Health States Value Function for MDP

State	Value
Critical (0)	155.52
Serious (1)	161.92
Stable (2)	165.39
Recovered (3)	168.99

Cite this article as:
Udok, U.V., Victor-Edema, U. A., & Ijomah, M.A. (2025). Stochastic game models based on markov decision processes for multi-agent decision making in clinical healthcare systems. *FNAS Journal of Mathematical Modeling and Numerical Simulation*, *2*(4), 31-44. https://doi.org/10.63561/jmns.v2i4.1121

Table 7 Treat and Wait Probabilities for Health State

State	Prob Wait	Prob Treat
Critical (0)	0.046	0.954
Serious (1)	0.035	0.965
Stable (2)	0.139	0.861
Recovered (3)	0.881	0.119

Table 8 Reward Function for Health State on Wait and Treat Actions $(\boldsymbol{R}(\boldsymbol{s},\boldsymbol{a}))$:

State s	R(s,Wait)	R(s,Treat)
0 (Critical)	153.50	156.53
1 (Serious)	159.62	162.93
2 (Stable)	164.57	166.39
3 (Recovered)	170.00	168.00

Discussion
Table 1 provides information for 20 (P001 to P020) patients states and a derived payoff or reward associated with that state / risk profile with the following column:
ID: Unique identifier for each patient. Severity: Categorical description of the patient's condition (Mild, Moderate, Severe). Risk: Categorical description of the patient's risk level (Low, Medium, High). Severity Score: A numerical score assigned to Severity (1 for Mild, 2 for Moderate, 3 for Severe). Risk Score: A numerical score assigned to Risk (1 for Low, 2 for Medium, 3 for High). The payoff is a numerical value that is derived from the Severity Score and Risk Score.

Table 2 shows the empirical frequencies of the "treat" simulation data. This table supports the numerical validity of table 1. It confirms that the treat transition probability matrix was empirically derived from simulation-based data. Hence, the transition matrix assumes treatment shifts probability mass towards better states, while reducing the poor health states.

Table 3 displays the reduced frequencies of the stimulated data for "Treat'' decision on table 3 by applying equation (1). The reduced data is obtained by dividing the frequency of the patient's state over the total number of patients. This illustrates that when critical patients are treated, 85% of them improve to serious state and 15% of them become stable and none recovers immediately. 75% of serious patients become stable and 25% recovers directly. Stable

Cite this article as:
Udok, U.V., Victor-Edema, U. A., & Ijomah, M.A. (2025). Stochastic game models based on markov decision processes for multi-agent decision making in clinical healthcare systems. *FNAS Journal of Mathematical Modeling and Numerical Simulation*, *2*(4), 31-44. https://doi.org/10.63561/jmns.v2i4.1121

patients largely transitioned to recovered (75%) confirming that treatment accelerated full recovery, recovered patients remain recovered (absorbing state). This analysis confirms that "treat" decision is highly effective

Table 4 demonstrates the empirical frequencies of the "wait" simulation counts. The high counts on the transition frequency from diagonal $(\text{critical} - \text{critical}) = 16$, $(\text{serious} - \text{serious}) = 10$. It implies that without treatment patients tend to persist in critical state or worsen from serious to critical states. The recovered state still acts as absorbing state. This transition matrix explains that with "wait" intervention, critical and serious state have a high chance of persistence or worsening while the stable state may slowly improve.

Table 5 displays the reduced frequencies of the stimulated data for "Wait'' action on table 4 by applying equation (3.1b). The reduced data is obtained by dividing the frequencies of the patient's states over the total number of patients in the simulation. Table 5 explains the reduced summary transition matrix for "wait" intervention where 80% of critical patients remained critical after "wait" action, 20% of serious patients became stable, and only 10% of stable patients recovered for wait action, while recovered patients remain in an absorbing state.

Table 6 explains the health state value function for MDP showing the values for each state (condition). For state 0 (critical state) is the lowest value, reflecting high immediate penalties, state (1): 161.92 displays a moderate value that indicates a better prospect than critical but still requiring "treat" action, state (2): 165.39 suggests natural recovery potential under with less urgency. This aligns interpretations like "likely stable" in low-risk severe cases, while recovered (3): 168.99 is the absorbing state that enhances full health with positive utilities.

Table 7 represents the optimal action probabilities derived from the policy of the MDP/POMDP model such that when both probability actions are added, it becomes unity .

For critical patients (state 0) implies that "treat" intervention in 95.4% of cases confirming that immediate intervention is optimal since "wait" decision carries a high mortality or deterioration risk or relapse. Serious (State 1) have a high probability of treatment (96.5%) consistent with the need of proactive care to prevent worsening. The recovered (state 3) describes that once recovery is achieved further treatment is usually unnecessary.

Table 8 displays the reward function for health states on wait and treat actions which gives the expected cumulative rewards for each action pair. For critical and serious states "treat" decision yields higher (e.g. 156.53 vs. 153.50) confirming that intervention improves outcomes in high-risk conditions. For stable state "treat" action remains marginally better (166.39 vs. 164.57) showing that proactive treatment still offers effective value

Conclusion

This study has successfully developed a stochastic game theoretic framework based on Markov Decision Processes for multi-agent decision-making in clinical healthcare settings. The framework has featured the patient health state (critical, serious, stable and recovered) agent's action being "treat" and "wait" decision.

The MDP components models the patient progression under the assumption of full observability of the patient's health state. The probability of transitioning from state s to state s' after taking action a was defined by a (4×4) matrix of $p(t)$ and $p(w)$. The immediate reward received for taking action a is defined by the value function representing the expected cumulative reward which was obtained by Bellman Optimality Equation.

Based on transition matrix in table 2, from a Critical state, a patient never stays Critical or recovers directly. They have a high probability of becoming Serious and a smaller chance of becoming Stable. This indicates effective intervention moving patients out of immediate critical danger. Also, from a Serious state, a patient either improves significantly to Stable or even recovers directly. This suggests the "Treat" action is quite effective here. Again, from a Stable state, a patient has a very high chance of recovering, or a smaller chance of remaining Stable. The recovered state is a very favorable state in a healthcare context in which "Recovered" is an absorbing state. Once a patient is recovered, stays recovered.

Cite this article as:
Udok, U.V., Victor-Edema, U. A., & Ijomah, M.A. (2025). Stochastic game models based on markov decision processes for multi-agent decision making in clinical healthcare systems. *FNAS Journal of Mathematical Modeling and Numerical Simulation, 2*(4), 31-44. https://doi.org/10.63561/jmns.v2i4.1121

Recommendations

The study therefore recommends that:

i. The healthcare agents and systems should improve clinical decision-making under uncertainty by applying Markov Decision Processes (MDPs) to minimize patient times spent in critical or serious health states, delays and costs of care in order to ensure evidenced-based support and overall system performance.

ii. The healthcare policy-makers should ensure that the developed-stochastic systems model patient health states and hospital resources together towards common goals in order to balance cooperation and competition, make care more affordable to benefit patients in speeding up their treatments and managing resources wisely.

iii. The framework's performance should be evaluated using simulation-based analysis to demonstrate quantitative and practical-oriented results, its ability to improve patient outcomes, optimize resource utilization, enhance tangible improvements in care and overall system efficiency in order to ensure practical and measurable outcomes.

References

Acuna, J. A., Zayas-Castro, J., Feijoo, F., Sankaranarayanan, S., Martinez, S. R. & Martinez, D. A. (2021). The Waiting Game; How cooperation between public and private hospitals can help reduce waiting lists. *Science Business Media*.

Adida, E., Maille, F., & Vial, J. P. (2018). A stochastic game approach to patient-physician interaction in chronic disease management. *Operations Research, 66*(5), 1223–1241. https://doi.org/10.1287/opre.2017.1704

Archetti, M. (2013). Evolutionary game theory of growth factor production: implications for tumour heterogeneity and resistance to therapies. *British Journal of Cancer, 109*, 1056–1062.

Arnetz, J. E., Winblad, U., Arnetz, B. B., & Höglund, A. T. (2007). Physicians' and nurses' perceptions of patient involvement in myocardial infarction care. *European Journal of Cardiovascular Nursing, 7*(2), 113–120. https://doi.org/10.1016/j.ejcnurse.2007.05.005

Bauch, C. T., & Earn, D. J. (2004). Vaccination and theory of games. https://doi.org/10.1073/pnas.0403823101

Basar, T. & Olsder, G. J. (1999). Dynamic Noncooperative Game Theory. *Society for Industrial and Applied Mathematics. 3, 10A*

Blake, A., & Carroll, B. T. (2016). Game theory and strategy in medical training. *Medical Education, 50*, 1094–1106.

De Bruijn, H., & Heuvelhof, E. T. (2018). *Management in networks: On multi-actor decision making* (2nd ed.). Routledge.

Filar, J. A. & Vrieze, O. J. (1997). Competitive Markov Decision Processes *Springer Science & Business Media.*

Folland, S., Goodman, A. C., & Stano, M. (2017). The Economics of Health and Health Care (8th ed. *Routledge.*

Hauskrecht, M. (2000). Value-function approximations for partially observable Markov decision processes. *Journal of Artificial Intelligence Research, 13*, 33–94.

Jennings, N. R., Sycara, K., & Wooldridge, M. (1998). A roadmap of agent research and development. *Autonomous Agents and Multi-Agent Systems, 1*(1), 7–38. https://doi.org/10.1023/A:1010090405266

Klein, G. A. (1993) A recognition-primed decision (RPD) model of rapid decision making. *Decision Making in Action: Models and Methods*

Komorowski, M., Celi, L. A., Badawi, O., Gordon, A. C., & Faisal, A. A. (2018). The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. *Nature Medicine, 24*(11), 1716–1720.

Liu, C., Liu, X., Wu, F., Xie, M., Feng, Y, & Hu, C. (2018). Using Artificial Intelligence (Watson for Oncology) for treatment recommendations amongst Chinese patients with lung cancer: Feasibility Study. *doi:10.2196/11087*

Long, D., Zheng, X., Lu, Q., & Zeng, Y. (2017). Optimal control for sequential treatment strategies of chronic diseases. *IEEE Transactions on Cybernetics, 47*(12), *4381-4392.*

Maskin, E., & Tirole., J. (2001). Markov perfect equilibrium. for surgeons and the operating room environment. *Journal of Economic Theory, 100*(2), 191–210.

McFadden, D., Kadry, B., & Souba, W. W. (2012). Game theory: Applications for surgeons and the operating room environment. *Surgery, 152*(5), 915–922.

Mendonça, F. V., Catalão-Lopes, M., Marinho, R. T., & Figueira, J. R. (2020). Improving medical decision-making with a management science game theory approach to liver transplantation. *Omega, 94*, 102050. https://doi.org/10.1016/j.omega.2019.102050.

Cite this article as:
Udok, U.V., Victor-Edema, U. A., & Ijomah, M.A. (2025). Stochastic game models based on markov decision processes for multi-agent decision making in clinical healthcare systems. *FNAS Journal of Mathematical Modeling and Numerical Simulation, 2*(4), 31-44. https://doi.org/10.63561/jmns.v2i4.1121

Mohr, D. C., & Bittar, H. G. (2014). Game theory analysis of medication adherence behavior. *Journal of Medical Systems, 38*(3), *9991*.

Myerson, R. (2004) Games with Incomplete Information Played by 'Bayesian' Players, I-III. Part I. *Management Science. DOI:*10.1287/mnsc.1040.0297

Osborne, M. J., & Rubinstein, A. (1994). *A course in game theory*. *MIT Press.*

Puterman, M. L. (2005). Markov decision processes: Discrete stochastic dynamic programming. *Wiley-Interscience.*

Reluga, T. C. (2011). Game theory of social distancing in a pandemic. Public library of science (*PLOS)*, *6*(8), e23640. https://doi.org/10.1371/journal.pone.0023640

Rubin, M., Pignatelli, C., & Haimowitz, I. (2016). Game theory and oncology: a systematic review. *Journal of Clinical Oncology*, *34*(15), *e20524-e20524.*

Scharpf, F. W. (1997). Games real actors play: Actor-centered institutionalism in policy research. *Westview Press.*

Schelling, T. C. (2010). A practitioner's approach. Economics and Philosophy, 26(1), 27-46.

Stubbings, L., Chaboyer, W., & McMurray, A. (2012). Nurses' use of situation awareness in decision-making: An integrative review. *Journal of Advanced Nursing*, *68*(7), 1443–1453. https://doi.org/10.1111/j.1365-2648.2012.05989.x

Vahdati, N., Vahdati, S. & Soltani, T. S. (2024). Design and Evaluation of Collaborative Decision-Making Application for Patient Care in the Emergency Department. *Health Science Report. 25, 7*(2)

Wu, M., Yang, M., Liu, L., & Ye, B. (2016). An investigation of factors influencing nurses' clinical decision-making skills. *Western Journal of Nursing Research*, *38*, 974–991. https://doi.org/10.1177/0193945916633458

Zhu, X., Jiang, L., Ye, M., Sun, L., Gragnoli, C., & Wu, R. (2016). Integrating evolutionary game theory into mechanistic genotype–phenotype mapping. *Trends in Genetics*, *32*, 256–268.

Cite this article as:
Udok, U.V., Victor-Edema, U. A., & Ijomah, M.A. (2025). Stochastic game models based on markov decision processes for multi-agent decision making in clinical healthcare systems. *FNAS Journal of Mathematical Modeling and Numerical Simulation*, *2*(4), 31-44. https://doi.org/10.63561/jmns.v2i4.1121