Proceedings of the 2025 Winter Simulation Conference
E. Azar, A. Djanatliev, A. Harper, C. Kogler, V. Ramamohan, A. Anagnostou, and S. J. E. Taylor, eds.

SIMULATING PATIENT-PROVIDER INTERACTION IN ICU ALARM RESPONSE: A HYBRID MODELING APPROACH

Zhaoqi Wang[1], Christopher Parshuram [2], and Michael Carter[1]

[1]Dept. of Mechanical and Industrial Engineering, University of Toronto, Toronto, ON, CANADA
[2]Dept. of Critical Care Medicine, The Hospital for Sick Children, Toronto, ON, CANADA

ABSTRACT

Patients in intensive care units (ICU) require continuous monitoring and care. When physiological abnormalities occur, an alarm is triggered to alert healthcare providers. The alarm response time is influenced by factors such as patient-population profile, care team configuration, staff workload, and unit layout. Response delays can directly impact patient outcomes, emphasizing the need for adequate emergency management capability. While prior simulation studies have explored ICU operations, they often oversimplify dynamic and concurrent interactions between patients and healthcare providers. This study presents a proof-of-concept hybrid simulation model that integrates discrete event simulation (DES) and agent-based simulation (ABS) to comprehensively represent a pediatric ICU environment. By simulating routine activities, patient-triggered alarms, and real-time interactions, the model investigates how varied resource configurations affect response times and outcomes. Built on scalable logic and realistic workflows, the model serves as a foundation for future clinical data integration and supports the development of ICU decision-support applications.

1 INTRODUCTION

In an intensive care unit (ICU), timely recognition and response to physiological abnormalities are critical to patient safety (Bridi et al. 2014). When a patient exhibits signs of distress, such as a drop in heart rate, blood pressure, or oxygen levels, an alarm is triggered to alert the healthcare team. However, responses to these alarms are not always immediate, and delays can result in adverse events such as worsened patient outcomes, prolonged hospital stays, or even death (Albanowski et al. 2023). The complexity of ICU settings, characterized by high-acuity patients and limited care team resources, makes timely responses challenging. Furthermore, the physical layout of the unit, including the distances between patient beds and nursing stations, further impacts the speed and efficiency with which care teams can respond to critical events (Obeidat et al. 2022). In practice, due to the high volume of alarms, regular alarms are often not well documented, and past studies have generally focused only on alarm counts, response times, or the impact of specific alarm types (Bridi et al. 2014). What is often overlooked is that alarms occur simultaneously with numerous other critical activities and patient care interventions within the ICU. The occurrence of these regular activities, including patient monitoring, drug-fluid therapy, and staff coordination (Marshall et al. 2017), creates a complex context that affects the timeliness of alarm responses.

This study investigates ICU alarm responses from care providers under various concurrent activities and care team configurations. We present a proof-of-concept hybrid simulation approach designed to capture both regular tasks and unplanned alarms. Scheduled activities and procedural alarm-handling workflows are modeled using discrete event simulation (DES). Meanwhile, patient-triggered alarms and staff working status, which dynamically change based on agent characteristics and current tasks, are modeled using agent-based simulation (ABS). Several unique features distinguish our work from existing ICU simulation studies (Williams et al. 2020; Ortiz-Barrios et al. 2023). First, the simulation incorporates shift changes, as research has shown that the workload and activities of the staff vary between different shifts (Debergh et al.

2012), which can influence availability and directly affect alarm response times. Second, the model allows one nurse to be responsible for one or two patients, as research indicates that variations in nurse-to-patient ratios directly affect patient survival rates (Lee et al. 2017). Finally, we designed customized performance metrics at both the population and individual agent levels, allowing users to assess the associations between care team configuration and alarm response performance. This granular analysis helps pinpoint bottlenecks and outliers, which would result in actionable insights for ICU resource planning and allocation.

2 LITERATURE REVIEW

Simulation has become an increasingly valuable decision support tool in healthcare due to its cost effectiveness and flexibility in scenario modeling (Laker et al. 2018). Among the most commonly used techniques, DES has been extensively applied to model time-based hospital processes, such as patient arrivals, diagnostic procedures, and treatment sequences. A 2021 review noted that emergency departments (EDs) are the most frequently modeled settings using DES, with approximately half of the studies focusing on improving time-sensitive metrics such as waiting times and throughput (Vázquez-Serrano et al. 2021). For example, Dosi et al. (2023) successfully implemented a DES model in a northern Italian ED to evaluate process improvement strategies, which led to the adoption of a pneumatic post system that reduced patient waiting times. In contrast, ABS simulates individual agents, such as patients, doctors, and nurses, interacting within a defined environment. This method is well-suited for modeling behaviors and phenomena such as disease transmission or patient behavior simulation (Silverman et al. 2015; Ajmal et al. 2024), which are difficult to represent using purely event-based methods.

Recognizing the strengths of each approach, researchers over the past decade have increasingly combined DES with ABS, and sometimes with machine learning, to build more comprehensive and realistic healthcare simulations. This hybrid approach has been used in various healthcare scenarios, such as emergency medical services and crisis management systems in hospitals (Anagnostou et al. 2013), integration of artificial intelligence (AI) for ICU demand forecasting (Ortiz-Barrios et al. 2023), and digital twin technology with DES for simulation of care provider and robot interactions (Anyene et al. 2024). Within intensive care settings, researchers have explored hybrid modeling for a range of applications, including the development of digital twins of the ICU using clinical data (Zhong et al. 2024) and the simulation of pathogen transmission during COVID-19 surges (Possik et al. 2022). These studies typically use DES to capture system-wide processes such as patient flow and resource allocation, while incorporating ABS to model individual behaviors and interactions. By combining the two methods, hybrid models provide valuable insights into both macro-level system operations and micro-level patient outcomes. Despite the growing interest in hybrid modeling for ICU applications, a notable gap remains: to the best of our knowledge, no existing studies have specifically applied hybrid simulation to model alarm response behavior. Therefore, we developed a proof-of-concept hybrid model using DES and ABS to simulate a hypothetical ICU environment and examine how care team configurations and task prioritization influence critical alarm response times.

3 METHODS

3.1 Modeling ICU Alarm Response

The simulation of the ICU alarm response is structured around three main components: input parameters, model environment design, and individual agent design. In this study, we chose not to incorporate real patient data at this stage. Instead, the model parameters are informed by publicly available research and expert opinions. The primary objective is to establish a foundational interaction model that remains adaptable and universal, regardless of the eventual data source or input type. By not incorporating real-world data, we eliminated potential noise and bias that could hinder the validation of the model output. This approach maximizes modeling flexibility and allows precise control over parameters to create desired theoretical conditions and scenarios. As a result, we can focus on refining the structure of the system and the behavioral logic of individual agents. The potential for integrating real clinical data is discussed in the Future Work section (Section 5.2). The following subsections elaborate on the input parameter designs and model logic.

3.1.1 Input Parameters

The overall model environment is specified through user-defined system parameters that represent resource demand and allocation in the ICU, including the Starting Shift, Patient Population Profile, Care Team Configuration, and Reference Dictionaries.

The Starting Shift parameter accounts for the dependence of the care team configuration and clinical activities on the shift schedules. Furthermore, emergency events have been observed to fluctuate according to patient acuity or clinical activity at different times of the day (Yartsev and Yang 2023). Referencing the same research, in the model, shifts are divided into three 8-hour windows: 00:00 to 08:00, 08:00 to 16:00, and 16:00 to 24:00. The user is required to select the starting shift at the initialization stage of the simulation, which determines the specific time period at which the simulation begins.

The Patient Population Profile is defined by three key factors: (1) the number of patients occupying the ICU, (2) their alarm frequency, and (3) the requirements of the single-bed ward. To simplify the modeling of resource demands, occupancy levels are used instead of patient arrival rates. Although this approach reduces randomness and makes the model less realistic, it provides greater flexibility in simulating specific patient census scenarios. For proof-of-concept modeling, we based our assumptions on on-site observations and assigned representative alarm frequencies ranging from one to five alarms per patient per hour, with higher values corresponding to the most acute single-bed ward cases and lower values reflecting low-acuity patients in shared-bed wards. In addition, patient ward requirements vary with clinical condition, ranging from non-infectious to infectious or immunosuppressed states. To simplify the simulation, we introduce the Single-bed Ward Requirement parameter, which represents demand for single-bed wards based on patient condition categories, with the upper limit constrained by ward availability.

For the Care Team Configuration, the simulation model includes two types of ICU staff: nurses and doctors. These roles were chosen because they have distinct activity patterns and are integral to both routine operations and emergency responses. Nurses serve as the primary caregivers in the ICU, providing continuous patient support. Their responsibilities include drug and fluid therapy, conducting diagnostic tests, and responding to alarms, among other tasks. According to literature reviews, many ICUs face resource constraints in which a single nurse may be responsible for the care of two patients simultaneously (McTavish and Blain 2024). To account for this variability, the model allows users to configure the nursing team by specifying how many nurses are assigned to one patient versus those assigned to two. Furthermore, on-site observations revealed that some nurses may experience idle periods during their shifts, either because they are not currently assigned to a patient or because their assigned patient has recently been discharged. In practice, the charge nurse is typically responsible for managing staff schedules and may reassign nurses as needed when new patients are admitted. To capture these dynamics, the model includes a parameter called the Number of Idle Nurses, which allows users to specify how many nurses are available but not actively assigned to specific patients. Beyond nursing staff, doctors also play a key role in ICU care, including leading ward rounds, responding to emergencies, and performing interventions. Their availability during a shift can have a significant impact on the speed with which emergencies are handled. Therefore, the model includes a parameter that lets users set the number of doctors available for each shift.

Two reference dictionaries, the Task Dictionary and the Bed Workspace Mapping Dictionary, are created for easy access to specific information during the simulation. The Task Dictionary includes both regular clinical activities and unplanned alarm events in the same format, each characterized by key attributes: location(s), duration(s), interruptibility, and the time required to disentangle. These task details are stored as strings and are parsed only when the task is about to be executed, helping to conserve computational resources during the simulation. The Bed Workspace Mapping Dictionary records the relationships between patient beds, rooms, and nursing stations, as nursing stations and monitoring areas are typically distributed according to the location of the wards. This dictionary allows staff agents to reference the appropriate locations when performing relevant clinical tasks.

3.1.2 The Model Environment Design

The model environment replicates the physical layout of the ICU and manages time-dependent events that repeat at the system level. The ICU floor is constructed using geometric shapes to represent key components such as patient rooms, nursing stations, beds, and connecting pathways. An essential time-based event in the simulation is the shift change, which occurs every eight hours to mirror the standard ICU staffing schedules (Yartsev and Yang 2023). During each transition, outgoing staff complete their tasks and perform a formal handover to the incoming team. Any alarm that occurs before the last second of the outgoing shift is addressed by the current nurse if they are idle; otherwise, it is handled by the incoming nurse. This ensures a smooth transition and uninterrupted patient care. In addition to simulating spatial and temporal dynamics, the main model environment also incorporates helper functions and global variables to support initialization and enable real-time tracking of the agent population. These include establishing an agent pool, producing dynamic patient lists for ward rounds, and detecting patients who require medical attention.

3.1.3 Individual Agent Design

In the simulation, the agents are divided into two categories: static resource agents and human agents. Bed spaces are modeled as static resources to track utilization and manage associations with other agents. Human agents, including patients, nurses, and doctors, are characterized by fixed parameters (e.g., walking speed), dynamic variables that evolve through interactions (e.g., number of interruptions), statecharts that define behavioral states, task lists assigned to each agent to manage scheduled duties (e.g., drug administration), where tasks are drawn from the predefined Task Dictionary with alarm-related tasks prioritized at the front of the queue and regular tasks added to the end, relationship mappings to track care responsibilities (e.g. nurse-to-patient assignments), and flowcharts that guide task execution and decision-making; the first five components align with ABS techniques, while the last one follows structured processes of DES.

Patient Agent Design: Each patient agent is assigned a Single-Bed Ward Requirement parameter, inherited from the overall Patient Population Profile. They also have a string parameter specifying their bed assignment and location, which is linked to a designated work area defined in the Bed Workspace Mapping Dictionary. In addition to these parameters, patient agents maintain variables that track their current alarm type, categorized as False Alarm, Assist Call, or Emergency. They also record current and average alarm response times as dynamic variables, which allows for real-time monitoring of staff responsiveness. Figure 1 presents the statechart that models the dynamic health status of each patient agent. Each patient transitions between two primary states: Stable and Alarm On. The Stable state indicates the patient does not require immediate medical attention. A transition to the Alarm On state occurs when an alarm is triggered according to the patient's predefined alarm frequency, which serves as an abstract surrogate for their acuity level.

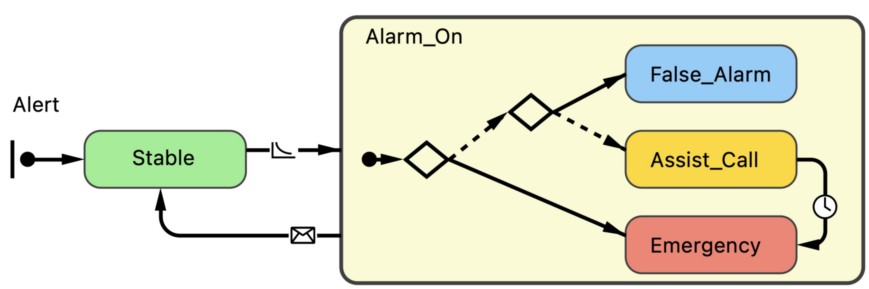

Figure 1: Statechart structure of patient agent alarm states and transitions.

In the Alarm On state, three sub-states (False Alarm, Assist Call, and Emergency) are designed to reflect the severity of alarm events, inspired by ICU emergency pattern research by Yartsev and Yang (2023). A False Alarm represents a non-urgent or technical issue that requires no intervention. An Assist Call signals a clinically relevant condition that requires prompt attention but is not immediately life-threatening. In contrast, the Emergency sub-state indicates a critical situation that demands urgent medical intervention. When the Alarm On state is triggered, the outcome is predetermined within the patient agent but remains unknown to the care team. Nurses are unaware of the alarm type until they respond, at which point the classification is revealed. This design reflects real-world ICU conditions, where providers must respond to alarms with limited knowledge of their urgency or severity.

Based on existing literature and expert opinion on ICU alarm types (Cho et al. 2016; Vreman et al. 2020), the default probability distribution for these alarm outcomes is set as follows: 80% for false alarms, 15% for assist calls and 5% for emergency alarms. Users may modify these probabilities to better align with specific simulation scenarios. Each patient agent maintains two key relational references to nursing staff: Primary Nurse and Current Nurse. The Primary Nurse is the nurse assigned to a specific patient for the duration of a shift. This assignment remains fixed throughout the shift and is only updated during shift transitions. The Primary Nurse acts as the primary caregiver and the point of contact for the patient, serving as a consistent reference for care responsibilities during the shift. In contrast, Current Nurse refers to the nurse actively involved in patient care at any given time. This relationship is dynamic and may temporarily differ from the Primary Nurse due to operational demands. For example, a nurse from the previous shift may be completing unfinished tasks or the Primary Nurse may be caring for another patient in cases where a nurse-to-patient ratio of 1:2 is implemented. This structure ensures an accurate representation of task delegation and continuity during handovers and shift changes.

The main flowchart for patient agents, illustrated in Figure 2, schedules regular care tasks every hour. These tasks are stored in the patient's task list, waiting to be retrieved and completed by the corresponding nurse. The receiving care delay block holds the patient for one hour, and any tasks not completed within that period are carried over to the next cycle. In addition, the flowchart includes a recovery check, which currently serves as a placeholder for future model enhancements, such as introducing a recovery variable to enable dynamic discharge policies based on patient recovery status.

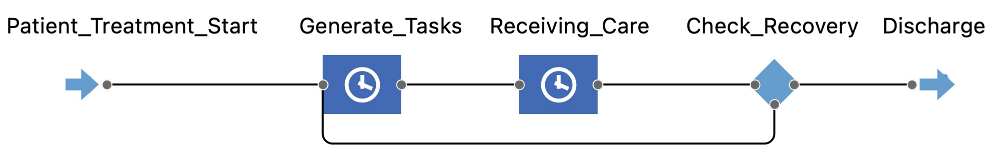

Figure 2: Task cycle flowchart for patient agent.

Nurse Agent Design: In the ICU, nurses are the primary caregivers and handle most patient-related tasks. In the model, these tasks are dynamically generated by patients, reflecting real-world workflows where care is driven by patient needs (Secunda and Kruser 2022). To ensure coherent simulation behavior, each nurse works exclusively with their assigned patient(s), and nurses without assignment remain idle (and do not assist others) until officially assigned. During simulation, nurses retrieve one task at a time from their patients, which helps prevent potential task synchronization issues. This design also simplifies shift changes, allowing nurses to be released after completing their current assignment, avoiding task reassignment.

For variables, each nurse agent has a shift variable that is used to assess overtime status and trigger release when the shift ends. They also maintain a comprehensive set of task-related variables that govern how tasks are loaded, interpreted, and executed. These variables include the originating patient, the ordered sequence of locations the nurse must visit to complete the task, and the corresponding duration required at each location. In addition, logical flags are also used to determine whether a task is regular, interruptible, or requires the nurse to accompany the patient. Each of these flags informs decision making within the agent's flowchart. This design enables the model to reflect realistic nursing workflows, including individualized task allocation, multi-stage task execution, and dynamic interruptions such as alarm events.

Nurse agents operate according to a working status statechart, shown in Figure 3, which monitors their current status and tracks utilization over time. A transition from Idle to Working occurs when the nurse receives a 'Work' message from the patient, which is typically triggered by the creation of a new task or the activation of an alarm that requires urgent attention. If the nurse is already in the Working state when an alarm is triggered and the current task is interruptible, the nurse switches to handling the alarm, while the original task is rescheduled. This mechanism allows nurse agents to remain responsive to the dynamic care demands of the ICU environment. In contrast, the transition from Working to Idle takes place once the nurse has completed all tasks for their assigned patients. At this point, the nurse enters an idle state, during which they may rest at the monitoring station or remain on standby until the next task or alarm arises. They also maintain a task list that tracks all tasks assigned to them. Under typical conditions, this list contains only one task at a time. However, in the event of an alarm, the alarm-related task is given top priority and placed at the front of the list. Any ongoing or routine task is deferred and rescheduled as the next task in sequence, allowing the nurse to prioritize the urgent response, which is in alignment with the real-world ICU practice.

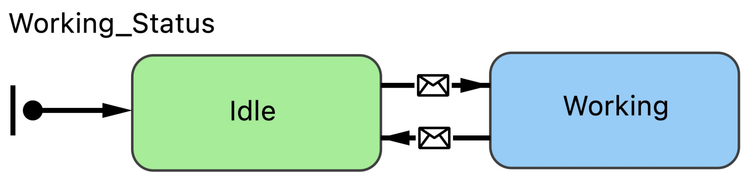

Figure 3: Nurse agent working-status statechart.

In addition to task management, nurse agents also have two patient relationship lists to simulate realistic care relationships in the ICU. The Primary Patient List includes patient(s) assigned to a nurse for the duration of a shift and remains static throughout that period, serving as a stable reference for care responsibility. In contrast, the Current Patient List is dynamically updated and reflects the patients to whom a nurse is actively caring at any given moment. This list evolves with task progression and is especially critical during shift changes, enabling seamless transitions and continuity of care. Together, these two lists ensure that the simulation captures both the scheduled and adaptive aspects of nurse-patient interactions in the ICU.

Two primary flowcharts govern the behavior of the nurse agent: one for regular care and the other for alarm response. These flowcharts model the decision-making processes and workflows of nurses in a typical ICU environment. The regular care flowchart, shown in Figure 4, outlines the steps that nurses follow to provide regular patient care. The process begins with an overtime check, where the model compares the nurse's assigned shift with the current simulation time. If the nurse is working overtime, they are redirected to complete shift change, where patient reassignment occurs, and the nurse is released from duty. If the nurse is not working overtime, the next block checks if the shift is nearing completion (less than 10 minutes remaining). In such cases, the nurse stops taking on regular tasks and responds only to alarms, if any occur. If the shift is not ending soon, the nurse then checks for unfinished tasks. If any remain, the nurse proceeds to the Load Task block to complete them. If no unfinished tasks are present, a new task is retrieved from an assigned patient. In the Load Task block, the task details are parsed and the task-related variables are set. The nurse then assesses the task type, determining whether it is a regular care activity or a non-typical task, and proceeds accordingly. Once the task is started, the nurse moves to the assigned location. During task execution, the system monitors for external interruptions, such as alarms. If an alarm occurs, the nurse transitions to the alarm response flowchart; otherwise, the task is carried out to completion. For tasks involving multiple locations, the nurse proceeds sequentially until all locations are covered. Once execution is finished, the nurse verifies that all task components are complete. If so, the nurse returns to the monitoring station to update records and retrieve the next task. This step sends the nurse back to the beginning of this flowchart. If no further tasks remain, the nurse enters a simplified idling flowchart, and their working status is updated to Idle.

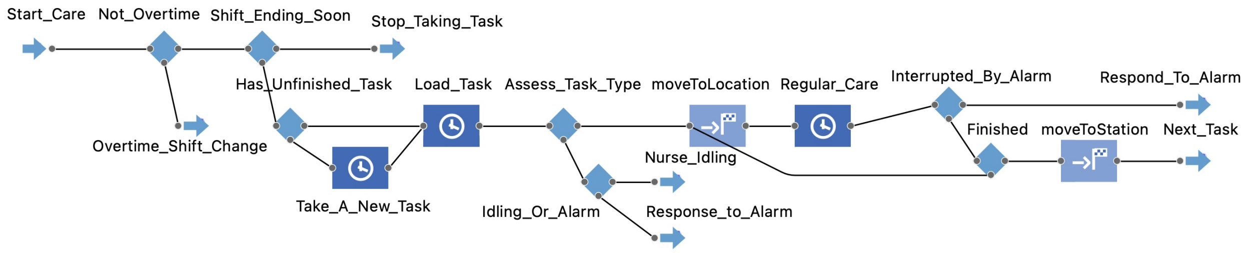

Figure 4: Nurse agent regular care flowchart.

When the patient triggers an alarm, the corresponding nurse immediately switches to the alarm response flowchart if they are not performing a non-interruptible task, illustrated in Figure 5. The nurse first moves to the patient's location and the alarm response time is recorded, capturing the interval between the alarm activation and the nurse's arrival. Upon arrival, the nurse assesses whether medical intervention is necessary. If the patient's condition requires the intervention of a doctor, the doctor is notified, and the nurse waits for the doctor to arrive to collaborate in managing the situation. If an escalation is not necessary, the nurse proceeds to handle the alarm independently, based on its type. For false alarms, the situation usually resolves in one minute. For actionable alarms, the resolution process generally takes between two and five minutes. After addressing the alarm, the nurse and the doctor, if involved, return to the monitoring station to document the event, review the patient's condition, and receive or provide further instructions before continuing with regular care activities.

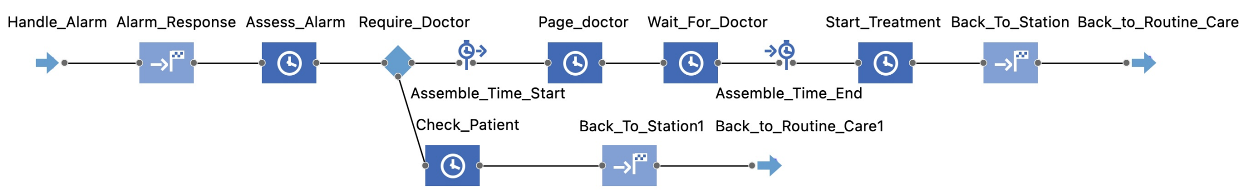

Figure 5: Nurse agent alarm response flowchart.

Doctor Agent Design: The doctor agent in the simulation is modeled with two primary responsibilities: conducting ward rounds and responding to emergencies. Structurally, the doctor agent shares the same components as the nurse agent, including identical parameters, statecharts, and task-related variables. This consistency standardizes performance tracking and agent design. In addition to these shared elements, the doctor agent maintains two lists to effectively manage patient care. The Current Patients List includes the patient the doctor is actively attending at any given time. During a typical ward round, this list contains only one patient. However, if an emergency occurs during a ward round, the list can temporarily include two patients, reflecting the dual responsibility until the emergency is resolved. The Critical Patients List tracks patients experiencing critical or emergency conditions. This list enables the doctor to prioritize clinical actions by maintaining visibility over patients requiring immediate intervention. The ward round flowchart, illustrated in Figure 6, begins by checking whether any patients remain to be seen. If so, the doctor selects one and initiates the ward round; otherwise, they proceed to the doctor agent duties flowchart. Once a patient is selected, a reception check confirms whether the patient has been successfully assigned to the doctor, preventing duplicate assignments. The doctor walks to the patient's bedside and performs the examination. The duration is modeled using a triangular distribution with a minimum of 4 minutes, a mode of 7 minutes, and a maximum of 10 minutes. These values were derived from on-site observations and published studies (Hillmann et al. 2022). This examination activity is preemptible. If a nurse issues an emergency page, the doctor abandons the ongoing check-up to attend to the urgent situation. Upon exiting the Check-Up block, the model determines whether the examination was completed or interrupted. If the actual duration is shorter than the expected duration, the task is marked as incomplete and rescheduled for a later time; otherwise, the doctor proceeds to the next patient. After all patients have been visited, a final check is performed to determine whether the ward round is complete. If no patients remain, the doctor transitions to other duties; otherwise, the process repeats.

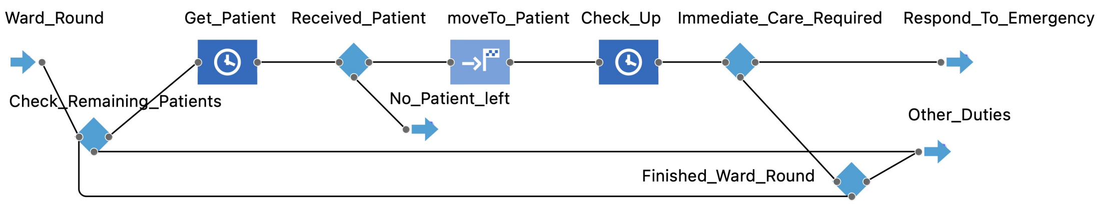

Figure 6: Doctor agent ward round flowchart.

The doctor agent duties flowchart, illustrated in Figure 7, models non-routine tasks performed by the doctor agent, including emergency response and administrative responsibilities such as documentation. Upon entry, the doctor first checks for any critical patients requiring immediate attention. If such cases exist, the doctor responds; otherwise, they proceed with administrative duties. When responding to an emergency, the doctor verifies whether a specific patient has been assigned. If so, they provide the necessary care; if not, the flow restarts. In the absence of active emergencies, the model checks whether the doctor has any remaining ward-round duties. If pending duties exist, the doctor transitions to the ward-round flowchart; otherwise, they proceed to assess their shift status. If the shift is ongoing, the doctor enters the Paperwork block, where interruptions may occur if a new task or emergency arises. If no further actions are required and the shift has ended, the doctor is released from duty. Unlike nurses, doctors are not tied to specific patients in a shift change. At the start of each shift, new doctor agents are generated and assigned to the doctor agent duties flowchart.

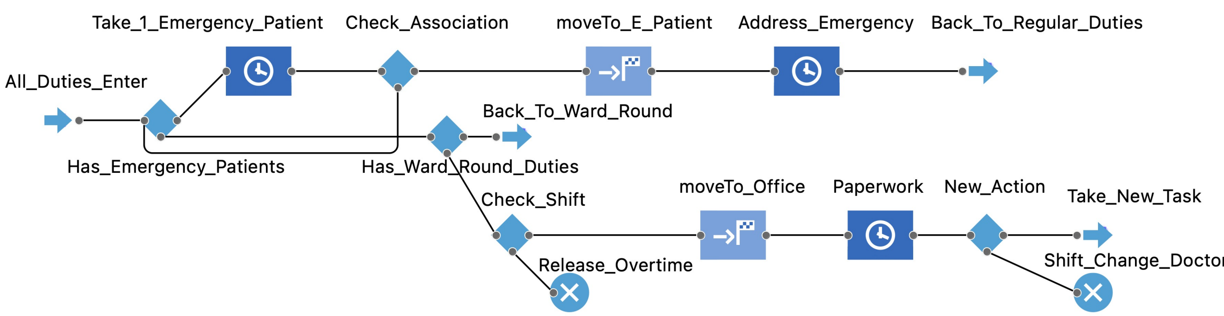

Figure 7: Doctor agent duties flowchart.

4 PROOF-OF-CONCEPT DEMONSTRATION

The ICU simulation model is designed to replicate the Pediatric Intensive Care Unit (PICU) at the Hospital for Sick Children (SickKids) in Toronto, Ontario, Canada. AnyLogic was selected as the simulation software because of its flexibility in supporting hybrid modeling. The software offers a Java-based object-oriented framework for logic customization. The spatial structure modeled in the simulation is based on the actual PICU floor plan to support realistic agent navigation and interaction. A demonstration version of the model can be accessed by clicking the following link: https://rb.gy/ub71ph. Users first configure the input parameters through an interface, after which patient, nurse, and doctor agents are initialized in the simulation environment. At this stage of development, the model serves as a proof of concept for a foundational ICU simulation framework, which can later be expanded with real-world data and routine care processes.

4.1 Design of Scenarios

We selected 23 bed spaces in the PICU (12 single-bed wards and 11 shared-bed wards) for simulation and tested three theoretical scenarios to evaluate how patient census and nurse-to-patient ratios affected alarm response times and staff utilization. In Scenario 1, 10 single-bed wards and 10 shared-bed wards were occupied (approximately 87% occupancy), and each of the 20 patients had a dedicated nurse. Scenario 2 maintained the same patient census but assigned 1:1 nursing only to single-bed patients, while shared-bed ward patients were cared for at a 1:2 nurse-to-patient ratio. Scenario 3 reduced the census to 6 single-bed patients and 5 shared-bed ward patients (approximately 50% occupancy), with all 11 patients receiving 1:1 nursing. For all three scenarios, doctor coverage was fixed at 1 doctor during the midnight shift, 3 doctors during the morning shift, and 2 doctors during the afternoon shift. Alarm response times and staff utilization metrics for each scenario were gathered and compared in the Results section.

4.2 Results

The hourly average alarm response time was calculated by aggregating the nurse alarm response times for each patient within the same hour, resulting in a unit-level performance indicator of overall responsiveness. Figure 8 compares these averages over a 24-hour window for all three scenarios. In Scenarios 1 and 3, where each nurse cared for a single patient, response times remained below 1 minute for approximately 85% of the day. In contrast, in Scenario 2, where half of the patients were managed under a 1:2 nurse-to-patient ratio, the average response time exceeded 1 minute for 90% of the period, peaking at nearly 4 minutes during the third shift. In all scenarios, response times showed brief increases after handovers at hours 8 and 16, reflecting delays caused by handover tasks and initial assessments as nurses settled into their duties. Furthermore, analysis of individual alarm response times showed that, for 90% of all patients, there were no differences greater than 1 minute across bed locations.

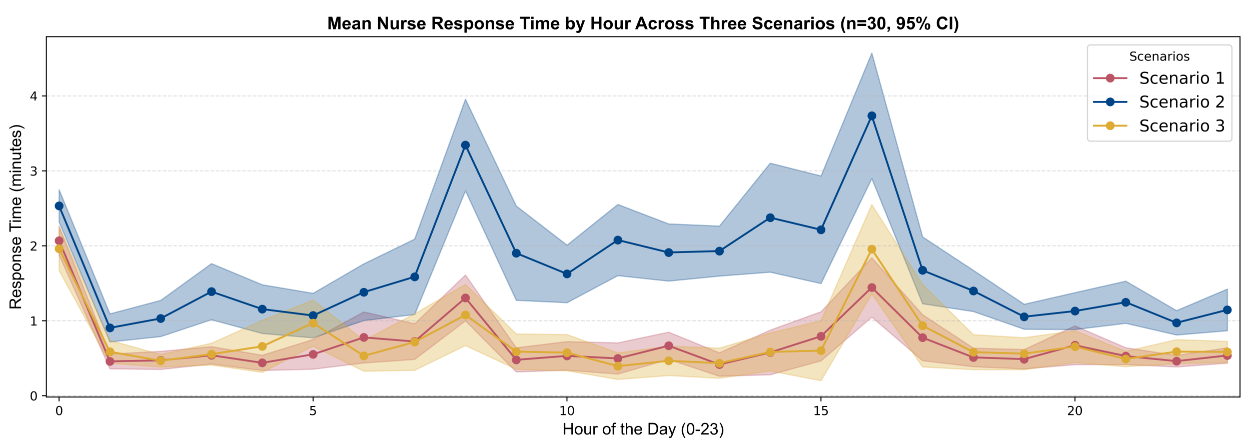

Figure 8: Nursing team average response time by hour across three scenarios.

Staff utilization was collected by surveying all staff's working status every 2 minutes and summarized as hourly averages. This metric reveals the general availability to respond to alarms for each type of staff. Figure 9 presents the utilization of the doctor team (left) and the nurse team (right) across the three scenarios over a 24-hour period. Doctor utilization varies early in the day but converges after hour 10, showing a similar trend across all three scenarios. No significant differences are observed between Scenarios 1 and 2, which have the same number of patients. Scenario 3, with fewer patients, maintains a consistently lower utilization throughout the period, ranging from 5% to 25% compared with the other scenarios. For nurses, utilization exhibits a distinct pattern: Scenario 2 consistently shows higher utilization, averaging approximately 61%, compared to 46% in Scenarios 1 and 3. This outcome aligns with expectations, as a 1:2 nurse-to-patient ratio naturally results in a higher workload compared to a 1:1 assignment.

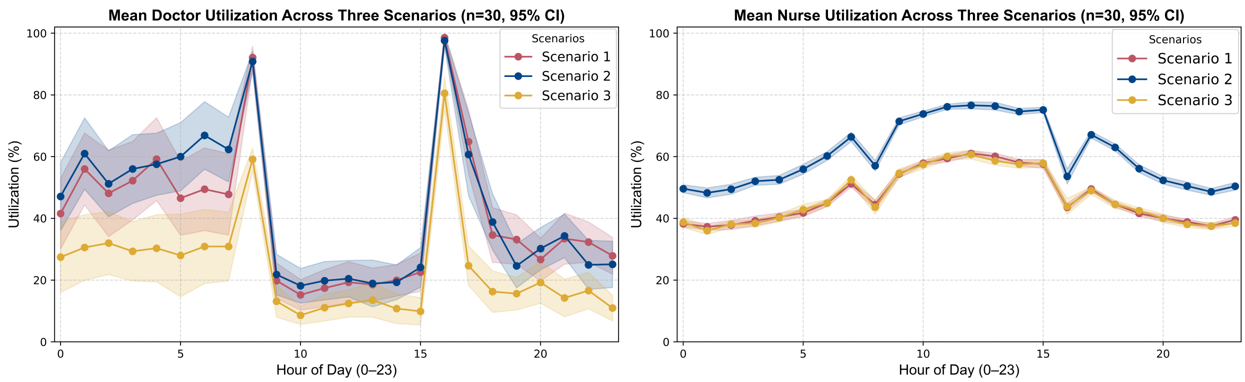

Figure 9: Doctor Utilization (Left) and Nurse Utilization (Right) across three scenarios.

The doctor's response time was measured for each emergency case. It was defined as the interval from when the nurse notified the doctor to when the doctor arrived at the patient's bedside. Figure 10 summarizes these times in the three scenarios. Scenario 3, characterized by the lowest patient census, demonstrated the best performance, with 90% of responses occurring within 5 minutes. With fewer concurrent emergencies competing for doctors' attention, queueing delays and backlogs were minimal. In contrast, Scenarios 1 and 2 showed longer right-hand tails. Their patient loads are approximately double those in Scenario 3, increasing the likelihood of overlapping emergencies competing for limited medical resources. As a result, the proportion of rapid responses decreased, while prolonged waits became more common, particularly in Scenario 2, where some nurses were responsible for two patients instead of one.

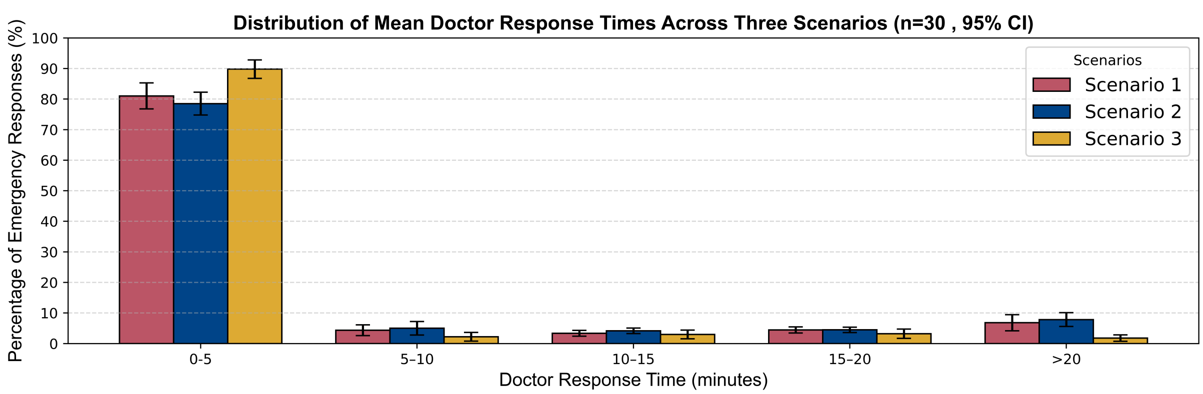

Figure 10: Distribution of doctor response times for emergency cases.

5 DISCUSSION

In the ICU, delays in responding to alarms can significantly influence patient outcomes. Variations in alarm types (e.g. false alarms, actionable alarms) and constraints on staffing (e.g., staffing ratios, competing priorities) make this problem hard to model and analyze. The simulation proposed in this paper provides a foundational model that captures the interactions between key ICU staff and simulates the logical workflows of each agent. Preliminary evaluations suggest that the results of the model are consistent with practical expectations. However, further development and validation are needed before fully integrating it into real-world ICU resource planning.

5.1 Limitations

As a proof-of-concept model, the simulation includes only patients, nurses, and doctors, omitting specialists such as respiratory therapists and pharmacists. In reality, ICU care is delivered through multidisciplinary collaboration and alarms related to ventilator management or medication safety often require specialized expertise. This exclusion underestimates the complexity and workload of coordination. Furthermore, the model also assumes a static patient census, overlooking admissions and discharges that affect staffing needs and alarm volume. In addition, patient acuity is approximated solely by alarm frequency. Although this approach enables basic differentiation between patients, it oversimplifies the complexity of real-world clinical conditions. Integrating data from the electronic health record (diagnosis and physiological scores) would provide a more realistic patient profile and improve the model's ability to predict patient outcomes.

5.2 Future Work

Future improvements to the ICU simulation model can be achieved by enhancing agent logic, incorporating various clinical roles, and integrating more comprehensive data. Admission rates could be introduced to simulate alarm response performance during peak periods or surges in patients. For discharges, a recovery policy could be used to track patient progress, and significant care delays could prolong hospital stays. This would provide insight into how the workload of the staff and the response times to alarms impact patient outcomes. Inspired by recent work from Ortiz-Barrios et al. (2023), future developments may explore a two-step simulation pipeline to improve patient demand modeling. In this approach, an AI model would first predict patient characteristics by learning from historical clinical records. These predicted profiles would then serve as input to the simulation, enabling a more data-driven and realistic representation of patient populations. This integration would greatly improve the fidelity of the model and support a smoother implementation of the simulation into existing hospital planning workflows and decision support systems.

5.3 Conclusion

The ICU simulation model developed in this study provides a foundational framework for analyzing care delivery dynamics, with a particular focus on alarm response under varying care-team configurations. The results are consistent with practical expectations and offer preliminary evidence of the influence of nurse-to-patient ratios on response times. However, further scenario testing and integration of real-world data are needed to strengthen the model's workflow logic and overall robustness. Despite its current limitations, the model serves as a starting point for future research aimed at improving the realism and applicability of ICU simulations for resource planning and operational decision making.

REFERENCES

Ajmal, H., E. Hunter, J. Duggan, C. Timoney, and C. Walsh. 2024. "Agent-Based Modelling of Mpox Infection Dynamics: Simulating Disease Transmission and Control Strategies". *Journal of Artificial Societies and Social Simulation* 27(2):6.

Albanowski, K., K. J. Burdick, C. P. Bonafide, R. Kleinpell, and J. J. Schlesinger. 2023. "Ten Years Later, Alarm Fatigue is Still a Safety Concern". *AACN Advanced Critical Care* 34(3):189–197.

Anagnostou, A., A. Nouman, and S. J. Taylor. 2013. "Distributed Hybrid Agent-Based Discrete Event Emergency Medical Services Simulation". In *2013 Winter Simulations Conference (WSC)*, 1625–1636 https://doi.org/10.1109/WSC.2013.6721545.

Anyene, G., C. Schultz, A. Nepomuceno, and I. Kim. 2024. "Digital-Twin Co-Simulation Framework to Support Informed Decision in Healthcare Planning and Management". *Simulation* 101(3):361–375.

Bridi, A. C., R. C. L. d. Silva, C. C. P. d. Farias, A. S. Franco, and V. d. L. Q. d. Santos. 2014. "Reaction Time of a Health Care Team to Monitoring Alarms in the Intensive Care Unit: Implications for the Safety of Seriously Ill Patients". *Revista Brasileira de Terapia Intensiva* 26(4):28–35.

Cho, O. M., H. Kim, Y. W. Lee, and I. Cho. 2016. "Clinical Alarms in Intensive Care Units: Perceived Obstacles of Alarm Management and Alarm Fatigue in Nurses". *Healthcare Informatics Research* 22(1):46–53.

Debergh, D. P., D. Myny, I. Van Herzeele, G. Van Maele, D. R. Miranda, and F. Colardyn. 2012. "Measuring the Nursing Workload per Shift in the ICU". *Intensive Care Medicine* 38:1438–1444.

Dosi, C., M. Iori, A. Kramer, and M. Vignoli. 2023. "Successful Implementation of Discrete Event Simulation: Integrating Design Thinking and Simulation Approach in an Emergency Department". *Production Planning & Control* 34(13):1233–1247.

Hillmann, B., D. Schwarzkopf, T. Manser, C. Waydhas, and R. Riessen. 2022. "Structure and Concept of ICU Rounds: the VIS-ITS Survey". *Medizinische Klinik-Intensivmedizin und Notfallmedizin* 117(4):276–282.

Laker, L. F., E. Torabi, D. J. France, C. M. Froehle, E. J. Goldlust, N. R. Hoot *et al*. 2018. "Understanding Emergency Care Delivery Through Computer Simulation Modeling". *Academic Emergency Medicine* 25(2):116–127.

Lee, A., Y. S. L. Cheung, G. M. Joynt, C. C. H. Leung, W.-T. Wong, and C. D. Gomersall. 2017. "Are High Nurse Workload/Staffing Ratios Associated With Decreased Survival in Critically Ill patients? A Cohort Study". *Annals of Intensive Care* 7(1):46.

Marshall, J. C., L. Bosco, N. K. Adhikari, B. Connolly, J. V. Diaz, T. Dorman *et al*. 2017. "What is an Intensive Care Unit? A Report of the Task Force of the World Federation of Societies of Intensive and Critical Care Medicine". *Journal of Critical Care* 37:270–276.

McTavish, C., and A. Blain. 2024. "Nurse-Patient Ratios: Current Evidence Report". Technical report, Canadian Federation of Nurses Unions, Ottawa, Ontario, Canada.

Obeidat, B., M. B. Younis, and E. Al-Shlool. 2022. "Investigations Into the Impact of Nursing Unit Layout on Critical Care Nurses". *Heliyon* 8(2):e08929.

Ortiz-Barrios, M., S. Arias-Fonseca, A. Ishizaka, M. Barbati, B. Avendaño-Collante, and E. Navarro-Jiménez. 2023. "Artificial Intelligence and Discrete-Event Simulation for Capacity Management of Intensive Care Units During the Covid-19 Pandemic: A Case Study". *Journal of Business Research* 160:113806.

Possik, J., A. Asgary, A. O. Solis, G. Zacharewicz, M. A. Shafiee, M. M. Najafabadi, *et al*. 2022. "An Agent-Based Modeling and Virtual Reality Application Using Distributed Simulation: Case of a COVID-19 Intensive Care Unit". *IEEE Transactions on Engineering Management* 70(8):2931–2943.

Secunda, K. E., and J. M. Kruser. 2022. "Patient-Centered and Family-Centered Care in the Intensive Care Unit". *Clinics in Chest Medicine* 43(3):539–550.

Silverman, B. G., N. Hanrahan, G. Bharathy, K. Gordon, and D. Johnson. 2015. "A Systems Approach to Healthcare: Agent-Based Modeling, Community Mental Health, and Population Well-Being". *Artificial Intelligence in Medicine* 63(2):61–71.

Vázquez-Serrano, J. I., R. E. Peimbert-García, and L. E. Cárdenas-Barrón. 2021. "Discrete-Event Simulation Modeling in Healthcare: A Comprehensive Review". *International Journal of Environmental Research and Public Health* 18(22):12262.

Vreman, J., L. M. van Loon, W. van den Biggelaar, J. G. van der Hoeven, J. Lemson, and M. van den Boogaard. 2020. "Contribution of Alarm Noise to Average Sound Pressure Levels in the ICU: An Observational Cross-sectional Study". *Intensive and Critical Care Nursing* 61:102901.

Williams, E., T. Szakmany, I. Spernaes, B. Muthuswamy, and P. Holborn. 2020. "Discrete-Event Simulation Modeling of Critical Care Flow: New Hospital, Old Challenges". *Critical Care Explorations* 2(9):e0174.

Yartsev, A., and F. Yang. 2023. "Patterns of Intensive Care Unit Emergencies". *Australian Health Review* 47(2):234–238.

Zhong, X., S. V. A. Neena, G. Y. Hou, Y. Dong, A. Lal, and O. Gajic. 2024. "A Hybrid Simulation Approach for Modeling Critical Care Delivery in ICU". In *2024 Winter Simulation Conference (WSC)*, 1233–1244 https://doi.org/10.1109/WSC63780.2024.10838861.

AUTHOR BIOGRAPHIES

ZHAOQI WANG is a Ph.D. student in Industrial Engineering at the University of Toronto, specializing in operations research and machine learning for clinical decision support and healthcare operations. His email address is mr.wang@mail.utoronto.ca.

CHRISTOPHER PARSHURAM is a Staff Physician and a Safety Scientist at the Hospital for Sick Children, with interest in the identification of significant clinical problems (medication safety, healthcare provider fatigue, cardiac arrest), their quantification, the design and development of system-level interventions, and the evaluation of these interventions as methods to improve patient outcomes. His email address is chris@sickkids.ca and his website is https://www.sickkids.ca/en/staff/p/christopher-parshuram/.

MICHAEL CARTER is a Professor in the Department of Mechanical and Industrial Engineering at the University of Toronto (since 1981) and the founding director of the Center for Healthcare Engineering. His research focuses on healthcare resource modeling. His email address is mike.carter@utoronto.ca and his website is https://che.utoronto.ca/professor-michael-w-carter/.