

SNFsim: A Discrete Event Simulator for Decision Support in Skilled Nursing Facilities

Caroline Strickland^{1*}, Brittin Wagner², Stanley Wang², Daniel J Lizotte^{1,3}

¹University of Western Ontario, London, Canada; ²PointClickCare, 5570 Explorer Drive, Mississauga, Ontario, Canada; ³Department of Epidemiology and Biostatistics, University of Western Ontario, London, Canada

Abstract We introduce SNFsim, an open-source discrete-event simulator for developing and evaluating reinforcement learning (RL) methods for multi-dimensional sequential decision support in Skilled Nursing Facilities (SNFs). SNFs play a vital role in the United States healthcare system, delivering specialized care to individuals with ongoing medical needs. Decision-making within SNFs is often complex due to their fast-paced and stochastic nature. SNFsim provides a modular and extendable simulation of major decision-making processes within SNFs, capturing many of the complexities and uncertainties existing in healthcare environments while still being flexible enough to allow for easy customization. Its potential uses are two-fold: First, as a test bed for the development and comparison of RL algorithms, and second, as the basis of a decision-support system that can be tailored to individual SNFs.

DOI: <https://doi.org/10.34196/ijm.00337>

1 Introduction

Reinforcement learning (RL) methods have been central to the development of data-driven decision-making in domains ranging from autonomous vehicle navigation (*Stafylopatis and Blekas, 1998*) to finance (*Hambly et al., 2023*) to video games (*Mnih et al., 2013*). In many domains, simulation environments that replicate real-world dynamics have been indispensable in allowing agents to learn policies without the risks often associated with direct real-world interaction. Widely used examples range from simpler systems like the cart-pole environment (*Nagendra et al., 2017*) to more complex platforms such as chess and shogi (*Silver et al., 2017*). However, there remains a notable gap in simulators designed specifically to support the exploration and development of RL methods tailored for multi-dimensional and multi-objective decision support, especially within the healthcare domain.

To help bridge this gap, we introduce the Skilled Nursing Facility Simulator (SNFsim), an agent-based discrete-event simulator that combines real-world health data with subject-area expertise to simulate decision-making dynamics in post-acute care facilities. Like microsimulation models, which are used to inform policy by predicting population-level outcomes under different scenarios (*Rutter et al., 2019; Spielauer, 2011*), SNFsim models individual patients with heterogeneous characteristics to predict facility-level outcomes under different operational policies. However, it incorporates elements more commonly found in discrete-event and agent-based models (*Karnon et al., 2012; Marshall et al., 2015*) such as resource constraints and feedback loops where facility-level decisions directly influence individual patient outcomes. This approach captures the dynamics and reward structures necessary to develop and test RL methodology capable of navigating multi-dimensional and multi-objective decision-making in resource-constrained healthcare settings, thus expanding the applicability of RL to real-world health services.

*For correspondence:
cstrick4@uwo.ca

©This article is distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

Author Keywords: healthcare operations, reinforcement learning, simulation, sequential decision support
© 2026, Strickland et al..

1. 1 Research Objectives and Contributions

Our primary goal is to provide a simulator and associated environment designed to support the development of RL methods for creating multi-objective decision support tools within SNFs. This environment, which is effectively an interface layer over the underlying simulator, should contain a sufficiently detailed state and action space, and capture multiple real-world outcomes of decision-making within SNFs.

Our main objectives for this work are to:

1. Provide a scalable open-source simulator calibrated using real-world datasets that is both modular and easily modified to fit user needs.
2. Develop a Gymnasium-compatible RL environment that exposes the simulators underlying dynamics through a structured interface.
3. Demonstrate the use of RL agents in learning policies for decision support within SNFs using the aforementioned environment.

By sharing our experiences in tackling the challenges of mathematizing decision-making within a system as intricate as SNFs, we aim to provide valuable insights for future efforts in creating simulation-based environments, particularly those designed for complex, multi-objective, and multi-dimensional decision-making processes in healthcare. Through the quantification of SNF operations, our work displays the complexities of healthcare management from a computational perspective.

1. 2 Related Work

Healthcare simulators have long played a crucial role in various aspects of healthcare delivery, training, and decision support. In the late 1960s and early 1970s, computational simulation modeling had emerged as a widely used technique for addressing hospital scheduling issues in the United States (*Robinson et al., 1968; Kolesar, 1970; Rising et al., 1973*). Following refinement of modeling techniques in the 1990s, the number of papers published describing simulation modeling in healthcare has grown substantially, extending to a wider range of settings and problems (*Günel and Pidd, 2010*). In using simulators, we reduce the need for direct interaction with the corresponding real-world systems. This allows for faster, safer, more cost-effective solutions to complex problems. In a 1978 survey of computer simulation in healthcare by England and Roberts (*England and Roberts, 1978*), the authors examine reports of 92 simulation models, offering insight into the extensive history of simulation within the healthcare domain.

Developing a simulator that models every aspect of a healthcare system is often impractical and infeasible (*Pidd, 1997*). Thus, many researchers assume a certain scope and level of abstraction when modeling these complex systems. Many simulators focus on a patient-level system existing within healthcare settings, for example, through Ahangers work in the computation, modeling, and simulation of the HIV-AIDS epidemic (*Ahanger, 2022*). In this work, a general mathematical model for the HIV-AIDS epidemics is introduced, specifically integrating vaccination strategies to understand the virus transmission dynamics in the population. Similarly, the work of Man et al. introduces a Type 1 diabetes simulator based on insights into insulin and glucose kinetics during hypoglycemia, accounting for nonlinear insulin-dependent utilization. This simulator provides a reliable framework for Insilico trials, testing glucose sensors, and validating closed-loop control systems (*Man et al., 2014*).

Simulators have also been used to develop and evaluate RL algorithms for health by providing a controlled environment to train, test, and refine decision-making strategies. The work of Jalalimanesh et al. introduces an agent-based simulator of vascular tumor growth based on collected biological data (*Jalalimanesh et al., 2017*). This research demonstrates the power of a simulation-based approach combined with RL for simulating and optimizing radiotherapy. Petersen et al. use biological simulation and deep RL to develop and evaluate personalized multi-cytokine therapies for the treatment of sepsis (*Petersen et al., 2018*).

Simulators that only consider facility-level decisions have also been developed. Sun et al. introduce a simulator that considers length-of-stay and discharge outcomes among nursing home residents with the goal of identifying promising staffing strategies (*Sun et al., 2023*). Butler et al. use discrete-event simulation for healthcare asset allocation through patient misplacement, where

patients are scheduled and assigned to the best alternative unit within a hospital due to a shortage of beds in the preferred area (Butler et al., 1992).

To the best of our knowledge, there has been no identified or documented simulator that models the consequences of accepting or denying referrals within SNFs while considering their relationship to facility-level decisions and their impact on multiple interrelated outcomes. Moreover, we are not aware of any simulators in the post-acute care domain that explore the interactions between multiple concurrently existing subprocesses and their reciprocal effects. SNFsim not only concentrates on the repercussions of referral decisions on various outcomes, but also examines the interdependence between referral decisions and staffing decisions, shedding light on their mutual influence and supporting the development of RL algorithms with multiple objectives in multi-dimensional action spaces.

2 Decision-Making in Skilled Nursing Facilities

In the following section, we provide an overview of decisions, states, and outcomes relevant to decision-making in SNFs, and we identify the components of these that are modeled by SNFsim.

2.1 Decisions

Decision-making in SNFs includes both patient-level decisions as well as facility-level decisions. Within these categories, SNFsim models patient referrals and staffing decisions, respectively. Figure 1 presents an overview of the referral intake flow and staffing decisions process within SNFs.

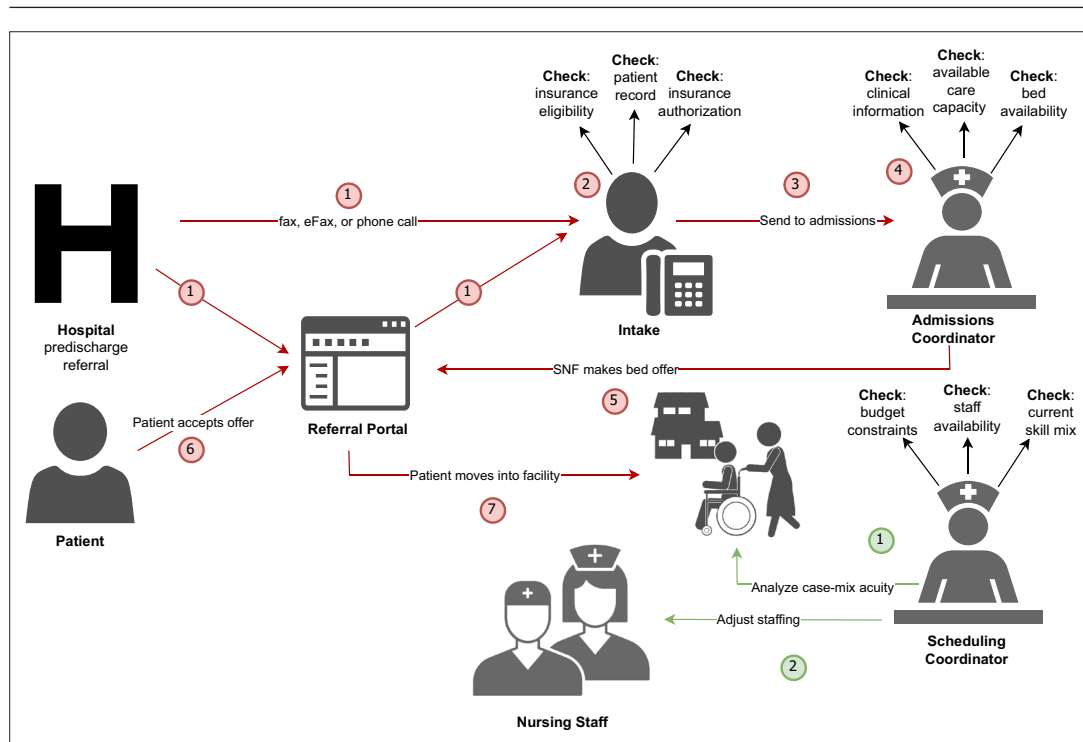


Figure 1. Skilled nursing intake flow and staffing process.

The red circled numbers denote the temporal ordering of the referral process, starting at the hospital and ending with a patient accepting an offer from a nursing facility. The green circled numbers denote the temporal ordering of a scheduling coordinator making staffing decisions. It is important to note that the intake flow and staffing process occur at different times.

2. 2 Patient-Level Decisions

The day-to-day activities in SNFs include many patient-level decisions, including the formulation of individualized treatment and nutrition plans based on patient status, discharge decisions, and advanced care planning. Patient-level decisions require a nuanced understanding of the patients health status and goals to provide comprehensive care.

SNFsim models the daily processing of individual patient *referrals* within SNFs. The referral intake process most often begins at the hospital via a pre-discharge referral, wherein the patients medical and financial details are communicated to a discharge planner within the hospital, either electronically or through traditional methods like phone or fax. This information is forwarded by the discharge planner to admissions coordinators at receiving SNFs, who are responsible for evaluating whether or not they can accommodate the care needs of the patient. Should the admissions coordinator determine that the facility is unable to offer appropriate care, the referral is declined and no offer is made. Otherwise, if the facility believes that it is able to meet the patients care requirements, the admissions coordinator extends an offer to the patient. Given that hospitals often dispatch multiple referrals for a single patient to various SNFs either concurrently or sequentially, the efficiency and promptness of the admissions coordinators response with an offer are crucial.

2. 3 Facility-Level Decisions

Facility-level decisions within SNFs often include things like financial decisions and policy implementation decisions. Facility-level decisions are responsible for shaping the strategic direction and overall quality of care in SNFs. These decisions require a more broad perspective that balances immediate needs with long-term objectives, ensuring that the facility remains a sustainable environment for long-term patient care.

SNFsim models staffing decisions within SNFs, which involves a careful evaluation of the needs of the current case-mix, budget constraints, and available nursing resources to ensure the delivery of high-quality care.¹ This process begins with assessing the acuity levels and specific care requirements of the resident case-mix, which dictate the necessary staff mix of nurses, certified nursing assistants (CNAs), and therapists. Staff scheduling is then optimized by managers to cover all shifts, taking into account fluctuations in the case-mix.

2. 4 State

We conceptualize the state of a SNF as all of the information required to make patient- and facility-level decisions described above. This includes information about the current case-mix, like demographics and care needs, as well as staff resources, financial metrics, and regulatory standing. Although each aspect of state holds its own significance, they are interconnected and together significantly influence the facilities overall effectiveness at making optimal patient- and facility-level decisions.

SNFsim maintains state information about the current case-mix (including detailed health information about each resident), facility-level financial information, and about current staffing levels. It also models transitions (moving from one state to another) in response to referral and staffing decisions.

2. 5 Outcomes

In the context of SNFs, an *outcome* is a measure of success or failure that reflects the effectiveness of the decisions being made within a facility. Outcomes in SNFs are important for understanding the overall performance of a facility, including clinical outcomes, financial outcomes, and operational outcomes. SNFsim supports the evaluation of four outcomes corresponding to each of these domains: rehospitalization, occupancy, patient reimbursement, and nursing costs.²

2. 5. 1 Rehospitalization

Rehospitalization is the process wherein patients, initially discharged from a hospital to a care facility for recovery, must return to the hospital due to complications or insufficient recovery progress. Rehospitalization events in a SNF can disrupt the stability of patients lives and elevate the

1. We use the term case-mix to describe a "mix of cases (patients)" receiving care.

2. This is often referred to as 'readmission'; in this work, we use the term 'rehospitalization' to make it more obvious that we are referring to the event where a patient leaves the SNF.

risk of medical errors associated with care coordination (*Mor et al., 2010*). Treatment in SNFs with historically low rehospitalization rates were found to causally reduce a [future] patients likelihood of rehospitalization (*Rahman et al., 2016*); thus, by minimizing rehospitalization, SNFs can improve patient outcomes and provide patient-centered care. Patient rehospitalization also represents a significant cost burden for SNFs, as it not only entails direct medical expenses but also impacts the performance metrics of the SNF. Thus, minimizing patient rehospitalizations ultimately leads to better patient care and improved operational functioning.

2. 5. 2 Occupancy

The occupancy of a SNF refers to the percentage of available beds that are currently occupied by patients, indicating the facility's usage rate and capacity utilization. In recent years, occupancy rates within SNFs in the United States have reached record lows (*Laes-Kushner, 2018*). Reduced occupancy introduces a variety of issues for SNFs, one of the most notable being that families seeking care for their loved ones often consider the occupancy rates of a facility as an indicator of its desirability and quality of care; low occupancy rates can raise concerns about the viability of a facility. Conversely, SNFs that are at or near maximum capacity face repercussions such as strained resources and potential impact on quality of care. As a general guideline, SNFs should strive to maintain an occupancy rate above 95.8% if possible, as facilities who do so are 2.09 times as likely to have greater efficiency scores (*Ozcan et al., 1998*). Higher (but not maximal) occupancies generally allow for the ability to accommodate slight fluctuations in demand, while still ensuring that the facility remains financially viable and able to deliver consistent high-quality care to its patients.

2. 5. 3 Patient Reimbursement

Each day, there exists some number of inpatients within any given SNF. These patients, typically through insurance, are responsible for reimbursing the facility for their care. The amount per day that a patient pays to the facility is directly linked to the complexity of their care needs, meaning that higher acuity implies a higher daily reimbursement amount from patient to facility.

In the context of SNFs, net revenue encompasses the total income earned from services provided to patients, including accommodation, medical care, and rehabilitation services. A portion of this net revenue is allocated towards operational expenses, including salaries for employees who deliver care, ensuring the facility can maintain a high standard of service. As the healthcare landscape continues to evolve, adequate net revenue is crucial for the long-term viability of SNFs. Simply, the more patients with high acuity (and therefore higher reimbursement rates) that a SNF accepts, the higher their net revenue becomes.

2. 5. 4 Nursing Costs

As mentioned previously, a portion of the net revenue of a SNF is allocated towards operational expenses. A major component of these expenses is compensating nursing staff, who are a central to delivering around-the-clock patient care. The less nurses that a SNF employs, the less they are required to pay out, thus increasing their net revenue. It is also true, however, that this would likely decrease the quality of care to patients within the facility, potentially increasing adverse outcomes like patient rehospitalizations.

3 The SNFsim Simulator

Developing a simulation model requires both formalizing what the model must encapsulate, as well as what can reasonably be excluded (*Tracy et al., 2018*). In the following, we describe how SNFsim is designed to capture the essential dynamics of patient flow, staffing patterns, and the effects of varying levels of patient care within a SNF setting. We begin by describing the data resources used to calibrate SNFsim and a simplified description of its transition dynamics, followed by a detailed description of how referrals are created and how the SNFs case-mix, staffing, and net revenue are updated as simulation progresses.

3. 1 Data Sources

We used two distinct primary datasets for calibrating SNFsim, each serving a unique purpose. The first dataset, which we will refer to as the *referral dataset*, is used to allow SNFsim to generate a realistic

distribution of referrals. The second dataset, referred to as the *classifier dataset*, is used to construct classifiers to decide whether or not a patient should be readmitted to the hospital on any given day during their stay.

Our referral dataset contains 643,770 referral records from 10,856 unique SNFs within 41 states in the United States between January 2020 and November 2022. Each data point captures detailed referral information including Length of Stay (LoS), primary diagnosis information, patient care needs, and insurance information.

Our classifier dataset contains 7948 observations from 1225 unique facilities between January 2020 and November 2022. Each data point captures information regarding the details of a patient's stay within a SNF. This includes LoS, mean number of certified nursing aide (CNA) hours per day during stay, patient care needs, primary diagnosis information, rehospitalization status, gender, and age.

3. 2 Design Philosophy

To accommodate a wide range of use cases, both customizability and simplicity were prioritized to ensure that the open-source codebase is easy to understand for as many users as possible. To achieve this, the Python programming language was chosen, and the RL environment API from OpenAI Gymnasium was used to implement the wrapper, reflecting Gymnasium's popularity within RL communities. **Figure 2** provides a simple example of how to use the Gymnasium environment.

```
// train.py

import gymnasium as gym
# Assuming PPOAgent is defined in custom_agents.py
from custom_agents import PPOAgent

# Load the custom environment
env = gym.make("snf\v0", render_mode="human")
agent = PPOAgent(env)

# Set the number of training episodes
num_episodes = 1e6

# Training loop
for episode in range(num_episodes):
    obs, _ = env.reset()

    while True:
        action = agent.select_action(obs)
        next_obs, reward, done, truncated, _ = env.step(action)

        agent.update_policy(obs, action, reward)

        obs = next_obs

        if done or truncated:
            break

env.close()
```

Figure 2. Code snippet demonstrating an example setup for training an RL agent using SNFsim.

3. 3 Overview of SNFsim Functionality

The transitions of SNFsim, broadly outlined in **Figure 3**, capture core operational dynamics and decision-making processes intrinsic to managing a SNF. This figure illustrates how various elements of action and state interact to produce the simulated dynamics of a SNF.

In the remainder of this section, we provide a closer look at the individual components, including the models used for predicting patient rehospitalizations, the generation of referrals at each timestep, and the formulation of outcomes. By providing these details, we aim to provide RL researchers with a comprehensive understanding of the intricacies involved in simulating the operations of a SNF and the potential of RL approaches to enhance decision-making within these environments. At the same time, we aim to provide health services researchers and providers with a clear understanding of what data and expert knowledge were used to specify the simulator.

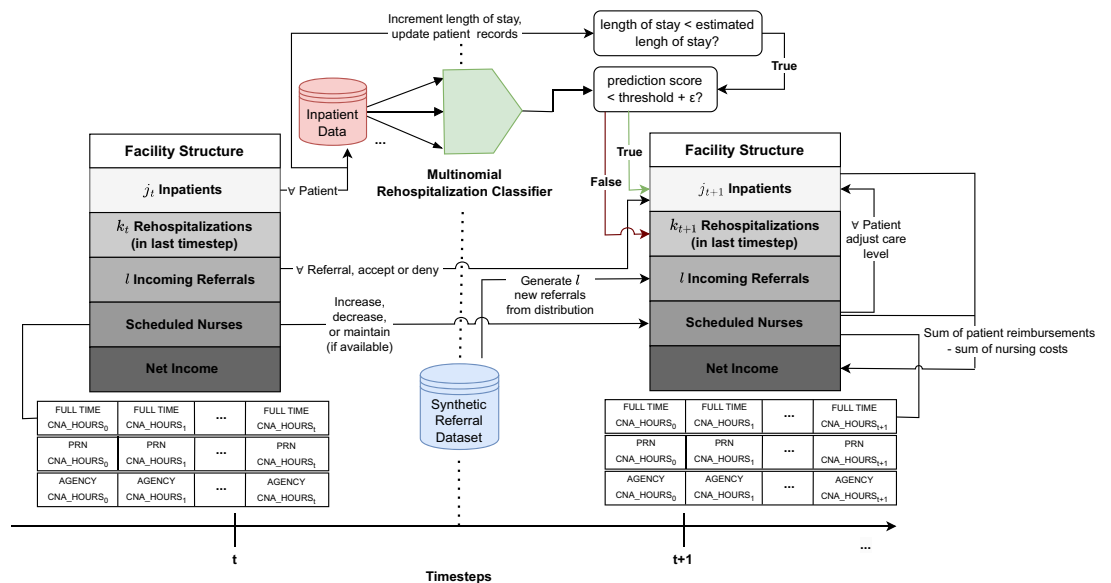


Figure 3. Simplified SNFsim simulation step.

We assume preset configuration values. Algorithm 1 presents an equivalent but more granular presentation of a simulation step.

3. 4 Referral Representation and Sampling

When generating a new referral, we are interested in sampling relevant patient features available at the point of referral. Typically, admissions coordinators receive the referral details outlined in **Figure 4** when a referral is being presented to a SNF. When generating referrals within SNFsim, we generate a subset of these referral features that actively contribute to the decision-making process for determining the acceptance or denial of a referral. Namely, we focus on a collection of clinical information and functional status, which serve as strong indicators of both the current health status and the financial implications associated with a referral.

We developed a synthetic dataset to be used within SNFsim. This approach allows us to generate realistic referral profiles by analyzing the statistical relationships within real-world data while ensuring complete patient privacy. By preserving the statistical structure of the original conditional distributions between features, we were able to generate representative synthetic patients that mirror the complexity and diversity of actual SNF populations without exposing personal health information.

To generate the synthetic dataset, we made use of the Synthetic Data Value (SDV) framework (**Patki et al., 2016**). Our original dataset, the referral dataset discussed in Section 3.1, was comprised of two datasets without patient-level identifiers. The first captured admission records containing ICD-10-CM codes, LoS, and insurance type. The second captured patient demographics containing age, gender, ICD-10-CM codes, and Payment Driven Patient Model (PDPM) classification codes.

To construct a unified dataset while preserving conditional distributions, we merged these datasets through diagnosis-based conditional sampling. For each admission record with a given IDC-10-CM

code, we randomly sampled a patient from the demographics dataset with the same diagnosis. This approach preserves the empirical conditional distribution $P(\text{demographics}|\text{diagnosis})$ observed in the original data, ensuring that patient characteristics associated with specific diagnoses remain realistic.

We then applied a Gaussian Copula Synthesizer to the merged dataset. This method separates marginal distributions from dependence structure (Sklar, 1959), allowing for the independent modeling of each variable's distribution while preserving multivariate relationships. The copula first transforms each variable to its empirical cumulative distribution function, applies inverse normal transformations, estimates the correlation matrix in the resulting multivariate normal space, and finally generates synthetic data by sampling from this multivariate normal distribution and applying inverse transformations. We configured the synthesizer with gamma distributions for LoS to capture right-skewed patterns, and beta distributions for age to respect bounded ranges from 18-120 years. Figure 5 shows age distribution by gender in the original and synthetic dataset (a) and LoS distribution by age group in the original and synthetic dataset (b).

We also implemented post-generation constraints. Since PDPM codes are diagnosis-dependent, we verified that each synthetic ICD-10-CM and PDPM code pairing had been observed in the original dataset. Invalid pairs (those never observed together) were corrected by resampling PDPM codes from the empirically valid set for that diagnosis.

Individual variable distributions were well-preserved, with an SDV column shapes score of 87.7%, indicating strong marginal distribution fidelity (detailed metrics in Table 1). Additionally, bivariate relationships across all 15 variable pairs showed good preservation with an SDV column pair trends score of 73.6%. Empirical validation of the trivariate distribution [age, LoS, gender] confirmed good multivariate fidelity with a variation distance of 0.173. However, complex non-linear interactions and categorical-specific patterns beyond pairwise correlations may not be fully captured. Further, clinical validity constraints were perfectly satisfied, with all of the unique 82,413 unique ICD-10-CM and PDPM code combinations in the synthetic dataset verified as observed in the original dataset, ensuring that all generated patient profiles represent clinically plausible (diagnosis, PDPM) pairings.

The synthetic dataset retained 92% coverage of original (diagnosis, PDPM) combinations while preserving 79% of diagnostic diversity (4,504 of 5,726 unique ICD-10-CM codes). The 1,222 excluded codes were overwhelmingly rare diagnoses: 99.4% (1,215 codes) occurred fewer than 10 times in the original data, collectively representing only 1.8% of total admissions. Notably, all top 500 most common diagnoses were fully retained, and these codes account for 85.7% of all admissions. This improves privacy protection by further preventing any reidentification through rare diagnosis combinations while still maintaining good coverage of clinically important scenarios.

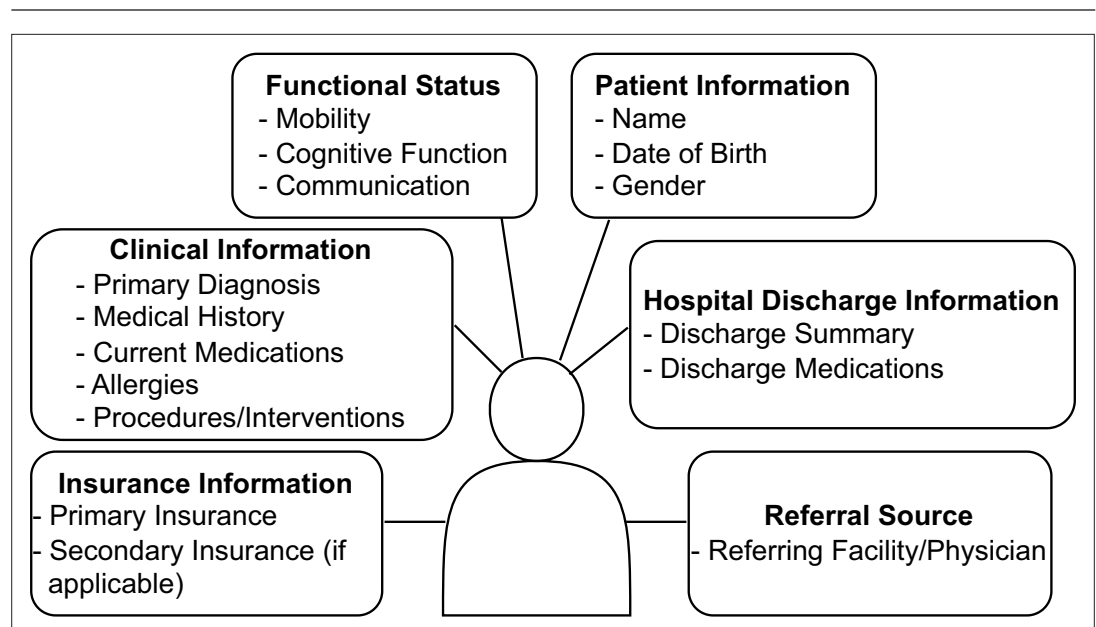


Figure 4. Patient information typically available to the SNF at the point of referral.

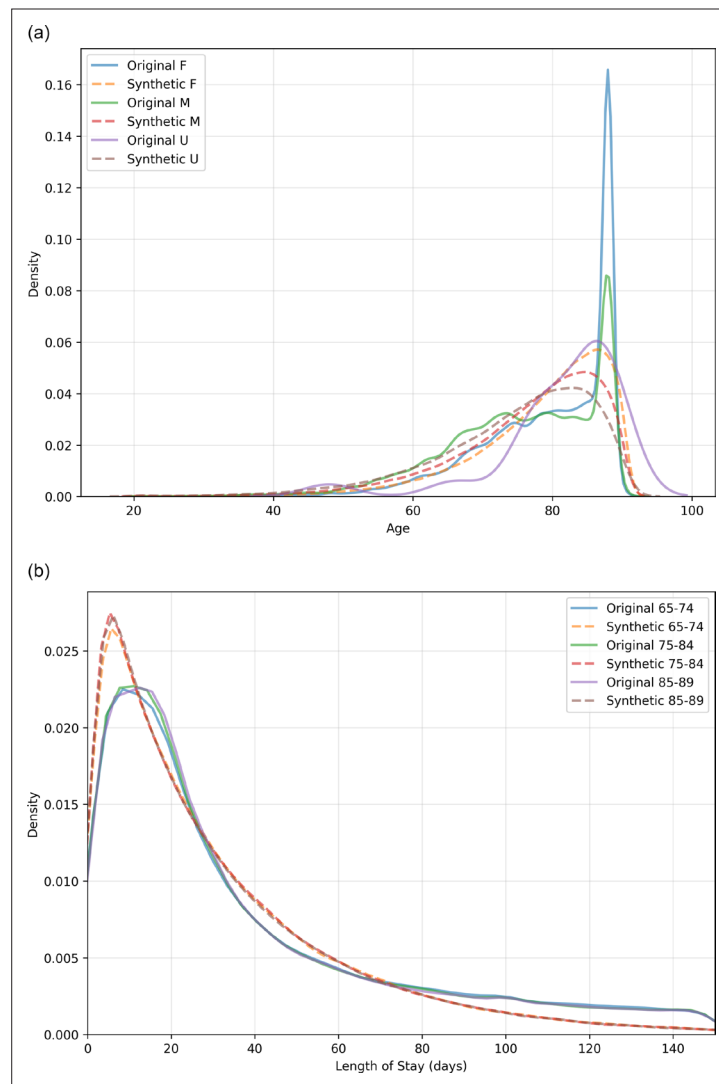


Figure 5. Bivariate relationships between a subset of columns in the generated dataset versus the original dataset. (a) Age distribution by gender. (b) LoS distribution by age group.

Table 1. Marginal similarity between original and synthetic data including the metrics mean diff (percentage difference in means), KDE Sim (kernel density overlap), TV dist (total variation distance), JS sim (Jensen-Shannon similarity), and cov (category coverage).

Variable	Mean Diff.	KDE Sim.	TV Dist.	JS Sim.	Cov.
<i>Continuous:</i>					
Age	0.02%	88.1%			
Length of Stay	13.8%	92.4%			
<i>Categorical:</i>					
ICD-10-CM			0.194	79.1%	78.7%
Insurance Type			0.150	83.6%	100%
PDPM Code			0.138	83.4%	98.2%
Gender			0.080	87.2%	100%

3. 4. 1 Patient Driven Payment Model

A core component in generating referral information lies in PDPM codes, which serve as a strong indicator of patient care complexity within SNFsim. We use PDPM codes in our synthetic data generation because they effectively encapsulate multiple dimensions of patient care needs and directly impact facility reimbursement, making them essential for realistic decision-making simulation. This choice is grounded in the PDPMs design to reflect the acuity and resource needs of the patient more accurately than previous models like the Resource Utilization Group (RUG-IV), which was therapy-driven. The PDPM adjusts payments based on the patients condition and care needs, rather than the volume of services provided (*Centers for Medicare and Medicaid Services*), making it a more patient-centered approach to reimbursement. A Health Insurance Prospective Payment System (HIPPS) PDPM code is comprised of five primary components: physical and occupational therapy (PT/OT), speech-language pathology (SLP), nursing, non-therapy ancillary (NTA) services, and an assessment type. Each of these five components has a corresponding Case-Mix Group (CMG) required for deciphering the corresponding character in the final code, as well as a Case-Mix Index (CMI) which is responsible for determining the daily reimbursement amount. **Figure A1** demonstrates the process of constructing a HIPPS PDPM code based on calculated patient care needs from each of the aforementioned components. The CMG of each of the five primary components is converted to a code value using **Tables A1 and A3**. The CMI and CMG can be calculated from **Table A4** (PT/OT component), **Table A5** (SLP component), **Table A6** (nursing component), **Table A11** (NTA component), and **Table A3** (5-day PPS MDS component). Once the CMG and CMI for each component is calculated, the 5-digit HIPPS PDPM code can be constructed. The reimbursement amount can then be calculated using Equation 1, which follows the official PDPM payment structure published by the Centers for Medicare and Medicaid Services (CMS), as defined in the PDPM Technical Report (**Acumen, 2018**).

$$\begin{aligned}
 \text{Daily Payment} = & \quad (\text{PT Base Rate} \times \text{PT/OT CMI}) \\
 & + (\text{SLP Base Rate} \times \text{SLP CMI}) \\
 & + (\text{Nursing Base Rate} \times \text{Nursing CMI}) \\
 & + (\text{NTA Base Rate} \times \text{NTA CMI}) \\
 & + (\text{Non-Case Mix Component}) \times \text{Adjustment Factor}
 \end{aligned} \tag{1}$$

It is important to note that the adjustment factor is calculated using **Table A10**, and dictates that the longer a patient is an inpatient within the facility, the less reimbursement the facility receives. Additionally, the base rate is calculated for each of PT/OT, SLP, nursing, and NTA using either **Table A8** or **Table A9**. Which table is used depends on whether the facility is located in a rural or urban area geographically. This is also where the non-case mix component is collected.

We developed an open-source module, *PyPDPM*, for calculating the precise reimbursement value based on a patients PDPM code and the number of days that patient has been a resident. For each timestep that a patient resides in the SNF, their calculated reimbursement amount is added to the facilitys overall net revenue. While real-world payment generally occurs monthly through a consolidated billing process (**Williamson, 2024**), SNF administrators know the daily per-diem rates and incorporate this information into admission and discharge decisions. Our per-timestep reward structure therefore captures the information available to decision-makers while also providing the immediate feedback necessary for effective RL training, avoiding the sparse rewards and difficult credit assignment that would result from monthly billing cycles.

3. 5 Facility Representation and Dynamics

SNFsim has a designated bed capacity, a limit on the number of staff (for each CNA option) available to attend to inpatients, and a list of n inpatients together with their relevant information at any given timestep.

The referral intake flow and staffing decision process within SNFsim are shown in **Figure 6**. It maintains the same sequential layout as that presented in **Figure 1**, but has abstracted away some of the real-world complexities and detailed procedural steps. In particular, by: assuming all generated referrals have valid insurance eligibility and authorization; assuming that if a referral is accepted then the patient accepts the offer; and removing the need for a referral portal by not keeping a log of unique patient IDs and their intake history.

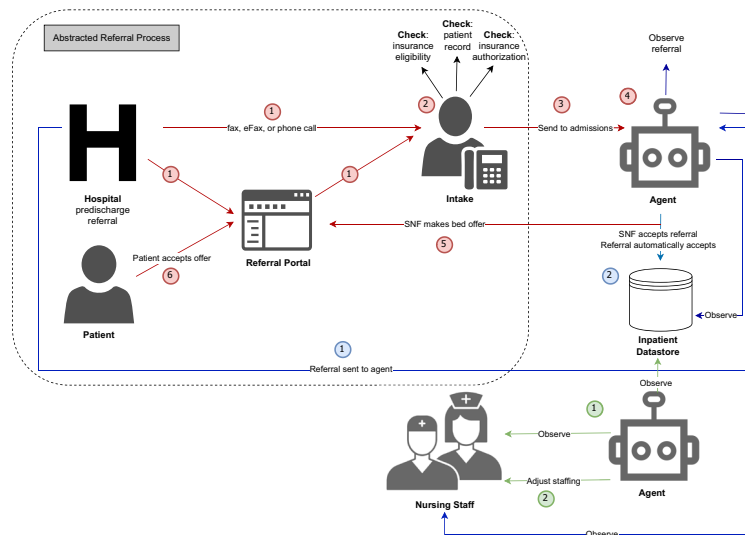


Figure 6. SNF intake flow and joint staffing decision process.

The red circled numbers denote the temporal ordering of a referral, beginning at the hospital and concluding with a patient accepting an offer from a nursing facility. This is abstracted, and is assumed rather than implemented within SNFsim. The blue circled numbers denote the temporal ordering of a referral within SNFsim, beginning with an empirically sampled referral and concluding with the SNF accepting the referral. The green circled numbers represent the temporal ordering of a simplified staffing decision process within SNFsim, beginning with assessment of the care needs of current inpatients and concluding with increasing, decreasing, or not altering current staffing hours.

The steps presented in Algorithm 1 comprise a single timestep of SNF sim at a high level, which we call a day.

Algorithm 1 Temporal Day in SNFsim (broad visual representation in Figure 3).

- 1: Input: State $S_t = \{P_t, O_t, H_t, C_t, R_t\}$, Action $a_t = \{A_t, \Delta_t\}$
- 2: Output: Next state $S_{t+1} = \{P_{t+1}, O_{t+1}, H_{t+1}, C_{t+1}, R_{t+1}\}$
- 3: where: P_t : patients, O_t : occupancy, H_t : nursing hours, C_t : finances, R_t : referrals
- 4: Initialize daily metrics: $Revenue_t \leftarrow 0$, $Rehospitalizations_t \leftarrow 0$, $Cost_t \leftarrow 0$
- 5: // Update staffing levels based on action
- 6: for each staff type $s \in \{FullTime, PRN, Agency\}$ do
- 7: $Staff_t^s \leftarrow \min(Staff_{t-1}^s + \Delta_t^s, MaxStaff^s)$ (Enforce staffing constraints)
- 8: $Cost_t \leftarrow Cost_t + \sum_{i=1}^3 Staff_t^s[i] \cdot Hours^s \cdot Rate^s$
- 9: end for
- 10: // Process referrals
- 11: $P_{new} \leftarrow$ Accept referrals based on A_t (0 or 1 for each referral in R_t) and bed availability B_{max}
- 12: // Calculate nursing hours per patient
- 13: $\vec{h}_t \leftarrow$ Nursing hours per patient per shift
- 14: $\alpha_t \leftarrow$ Staffing modifier based on ratio of actual to expected care
- 15: // Update patients and compute rehospitalizations
- 16: $P_{remove} \leftarrow \emptyset$
- 17: for each patient $p \in P_t$ do
- 18: Update patient day count, accumulate revenue, and record nursing hours received
- 19: Calculate shift-specific understaffing penalties $\vec{\epsilon}$
- 20: $p.r \leftarrow$ Rehospitalization risk based on patient data, staffing, and care quality
- 21: if $p.r > threshold$ then
- 22: Mark patient for rehospitalization and update count
- 23: end if
- 24: end for
- 25: Update patient population: $P_{t+1} \leftarrow (P_t \setminus P_{remove}) \cup P_{new}$
- 26: Update occupancy:
- 27: $O_{t+1} \leftarrow \frac{|P_{t+1}|}{B_{max}}$ {Occupancy as proportion of available beds}
- 28: Update finances:
- 29: $C_{t+1} \leftarrow C_t + Revenue_t - Cost_t$
- 30: Set next-day nursing hours:
- 31: $H_{t+1} \leftarrow H_t$ {Staffing decision at time t determines hours for day $t + 1$ }
- 32: Generate new referrals for next day: R_{t+1}
- 33: Return $S_{t+1} = \{P_{t+1}, O_{t+1}, H_{t+1}, C_{t+1}, R_{t+1}\}$

3. 5. 1 Rehospitalization model

SNFsim is designed in part to reflect the impact of multiple different factors on patient rehospitalization outcomes. To ensure optimal patient care, SNFs must consider both staffing levels and patient characteristics that influence rehospitalization risk. Based on consultation with healthcare experts and discoveries in our previous work (*Strickland et al., 2023*), we identified several key variables that impact patient rehospitalization.

To quantify how individual factors influence rehospitalization risk, we conducted multivariate logistic regression analyses using patient data from 7,948 observations (588 hospitalizations, 7.4% of the dataset). **Table 2** presents four model specifications. Model A isolates patient demographic factors (age and reimbursement) to demonstrate reimbursements expected positive association with readmission (as one would assume that cases with more complex care needs would have a higher likelihood of rehospitalization). As assumed, higher case complexity increases rehospitalization risk (coefficient = 0.188, $p < 0.001$). Model B includes LoS and key therapy indicators with minimal multicollinearity with a Variance Inflation Factor (VIF) < 3. Model C incorporates eight core clinical features, although multicollinearity is present (VIF > 10 for several features). Notably, in this model, the coefficient sign for patient reimbursement (a proxy for care needs) switches from positive (like in model A) to negative. Model D extends to 13 features total by adding categorical operational factors (diagnosis ICD-10-CM chapter, facility, payer type) as well as other relevant variables. These categorical features contribute to prediction but are excluded from the table due to space constraints. Notably, adding these operational features to logistic regression slightly reduces performance (AUC falls from 0.873 in model C to 0.867 in model D), showing that linear models do not benefit from these added complex interactions.

For optional use in SNFsim, we implemented a Random Forest classifier using all 13 features from Model D. Random Forest is well-suited for this prediction task as it naturally captures non-linear interactions between facility and patient characteristics. Ultimately, while 13-feature logistic regression achieves an AUC of 0.867, Random Forest achieves an AUC-ROC of 0.921 and an AUC-PRC of 0.592 (**Figure 7**). The higher AUC-ROC in this case indicates better overall ranking of patients by risk, while the AUC-PRC (though modest) represents an 8.0-fold improvement over the baseline rate of 0.074, demonstrating strong predictive performance given the substantial class imbalance typical in rehospitalization datasets and comparable in performance with similar classification models (*Pauly et al., 2019; Lou et al., 2025; Chandra et al., 2019*).

Note that in SNFsim, whether or not a patient is rehospitalized in a given timestep is not necessarily deterministic based on model output. The models predicted probability is adjusted by staffing-based multipliers (configurable) and noise and compared to a threshold (readmission_threshold in **Table 3**). If this threshold is exceeded, the patient is rehospitalized. This adds additional stochasticity and incorporates incentive for balanced nursing hours across shifts.

Table 2. Multivariate Logistic Regression Results

Feature	Model A		Model B		Model C		Model D	
	Coef.	P	Coef.	P	Coef.	P	Coef.	P
LoS			-3.173	***	-3.180	***	-3.193	***
meanHours					-0.216	***	-0.215	***
reimbursement	+0.188	***			-0.344	*	-0.329	*
age	-0.028				-0.005		-0.005	
NPG			-0.230	***	-0.274	***	-0.293	***
SLP			+0.077		+0.100		+0.083	
PT_OT					+0.033		+0.035	
NTA					_0.448	**	_0.428	**
days_since_start							+0.052	
gender							+0.017	

Additional categorical features in Model D (not included):

	Model A	Model B	Model C	Model D
ICD-10-CM Chapter (***) , facility, insurance type				
Number of features	2	3	8	13
AUC-ROC	0.550	0.867	0.873	0.867
Max VIF	14.7	2.1	33.3	36.6

Model D contains full feature set used in Random Forest (RF AUC=0.921 vs. Logistic Regression AUC=0.867). ***p < 0.001, **p < 0.01, *p < 0.05. N=7,948; 588 (7.4%) readmissions.

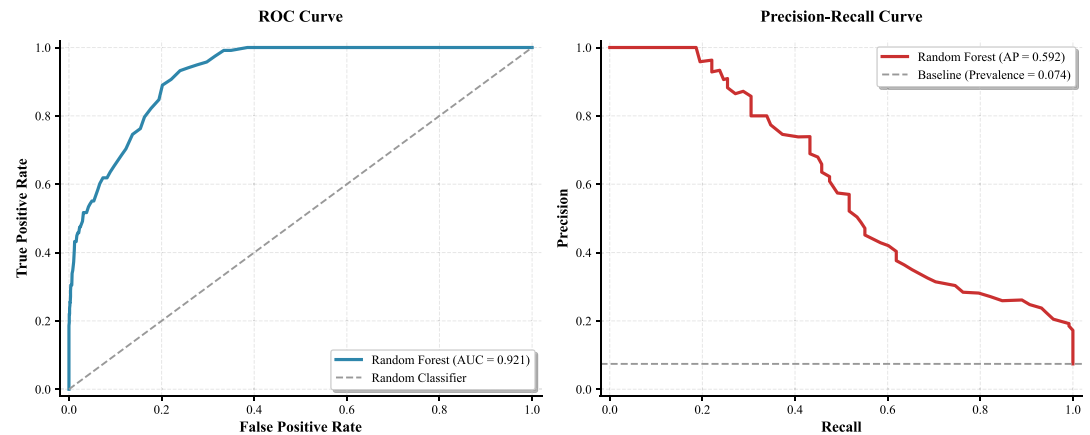


Figure 7. Random Forest model performance for patient rehospitalization prediction.

ROC curve (AUC = 0.921) shows excellent discrimination between rehospitalized and non-rehospitalized patients. Precision-Recall curve (average precision = 0.592) demonstrates strong performance relative to baseline prevalence of 7.4%. The model substantially outperforms logistic regression (AUC = 0.867) using the same feature set.

3. 5. 2 Tailoring SNFsim

Prior to simulation, users must configure key parameters of the SNF to match set up their preferred environment. **Table 3** presents the required configuration parameters along with their default values. These parameters include facility characteristics such as the total number of beds and optimal occupancy targets, staffing resources including the maximum available numbers of full-time, PRN (pro re nata or as needed), and agency CNAs, as well as operational factors like daily referral rates and minimum required nursing hours per patient day.

The selection of the facility's state location is important for accurately simulating financial aspects of operations, as each state within the United States has different hourly wage rates for nurses (of each subcategory). This selection impacts staffing costs and, consequently, affects the overall budget management and profitability of the SNF. The current state-specific wage data used in the simulation can be found in **Table A12**.

Table 3. Required SNF Configuration Parameters

Parameter	Description	Default Value
total_beds	Number of beds in the facility	100
occupancy_bounds	Target min/max occupancy range	75%/90%
nursing_hours_target	Target min/max nursing hours per patient	2.5/3.5
full_time_cna	Total available full-time nursing assistants	20
prn_cna	Total available pro re nata (as-needed) staff	10
agency_cna	Total available agency-contracted staff	15
referral_rate	Mean number of daily referrals	5

Parameter	Description	Default Value
min_cna_hours	Mandated nursing hours per patient day	2.8
fac_state	State where the SNF is located	New York
readmission_threshold	Threshold for readmission prediction	0.7

3. 6 OpenAI Gymnasium Environment

In this section, we document a custom OpenAI Gymnasium environment, `snf_v0`, which serves as a practical example of how users can create custom-tailored simulation environments based on SNFsim that can be easily used as testbeds for RL method development and application.

The environment described in this section is a direct operationalization of the facility representation introduced in Section 3.5. Specifically, the patient population, staffing configuration, occupancy levels, financial metrics, and referral generation defined in Section 3.5 make up the elements of the state space in `snf_v0`, while the referral acceptance and staffing adjustment steps correspond to the environments action space. Thus, this section translates the conceptual facility dynamics into a formal multi-objective Markov Decision Process (MO-MDP) compatible with RL.

3. 6. 1 State Space

The underlying simulator maintains a full Markovian state representation, $S_t = \{P_t, O_t, H_t, C_t, R_t\}$, which includes the complete set of inpatients and their associated attributes, facility occupancy, available nursing hours, cumulative financial metrics, and the full set of generated referrals. This state corresponds directly to the facility representation described in detail in Section 3.5 and is used internally by SNFsim to compute transitions and patient outcomes.

Because many patient-level features are high-dimensional, the environment does not expose the entire state to the agent. Instead, the agent receives a condensed observation, described in Section 3.6.2, that solely captures the operational information required for making decisions. Simply, `snf_v0` is a partially observable MOMDP in which $\mathcal{O} = f(S_t)$. The observation space provides a structured, lower-dimensional subset of the full simulator state, S_t .

3. 6. 2 Observation Space

While the full internal state, S_t , contains all patient-level and facility-level information, the observation space provides a structured subset of this internal information that is relevant for decision-making and avoids incorporating high-dimensional variables that may not be necessary for effective decision-making. Thus, the observation space is not identical to the state space, and the environment is partially observable. Our composite observation space includes three components; a condensed facility state vector, a staffing vector, and a fixed-size matrix of referral features. Let $\mathcal{O} = (f_t, s_t, r_t)$, where f_t summarizes the facility's operational status, and is formalized as:

$$f_t = [O_t, \text{Rev}_t, \text{Cost}_t, \bar{r}, N_t], \quad (2)$$

with O_t denoting occupancy rate, Rev_t the total daily reimbursement, Cost_t the daily staffing costs, \bar{r} the mean rehospitalization risk across all current inpatients, and N_t the number of residents currently admitted.

The staffing vector s_t , the second element of \mathcal{O} , encodes the total CNA hours scheduled across three shifts including day, evening, and night. This can be written as:

$$s_t = [H_t^{\text{day}}, H_t^{\text{evening}}, H_t^{\text{night}}]. \quad (3)$$

Finally, the referral matrix r_t provides a compact description of n candidate referrals at time t , with each row containing the expected daily reimbursement, expected LoS, and age of a prospective admission.

3. 6. 3 Action Space

In the `snf_v0` environment, the range of actions an agent can perform at some timestep t is defined by its action space. For this environment, we use a multi-component action space to accommodate the detailed decision-making required. This action space consists of both referral acceptance decisions and staffing adjustments. For referral decisions, we implement a MultiBinary space of size `referral_rate`, allowing the agent to make binary accept/reject decisions for up to `referral_rate` potential admissions simultaneously. Second, for staffing management, we employ nested Box spaces that reflect different operational realities for each staff type. Full-time staffing adjustments range from $[-2, 2]$ CNAs per shift, representing incremental changes to a persistent staff. In contrast, PRN and agency staffing use absolute values from $[0, \text{max}]$ for each shift, where `max` is the total available staff of that type. This staffing model mirrors real-world operations where PRN and agency staff are scheduled daily as needed, with staffing levels resetting to zero at the beginning of each day, while full-time staff schedules persist with incremental adjustments between days.

3. 6. 4 Reward Signal

The reward signal defines the objective measure to be optimized by the learning process. We construct a vector-based reward function that captures key operational objectives.

Occupancy Reward

The occupancy reward, R_{occ} , evaluates how well the facility maintains an optimal census level, normalized to a range of $[-1, 1]$:

$$R_{\text{occ}}(O_t, T_l, T_u) = \begin{cases} 1 - 2 \cdot \frac{O_t - T_u}{1 - T_u}, & \text{if } O_t > T_u \\ -1 + 2 \cdot \frac{O_t}{T_l}, & \text{if } O_t < T_l \\ 1, & \text{if } T_l \leq O_t \leq T_u, \end{cases} \quad (4)$$

where O_t represents the occupancy rate at time t , and $[T_l, T_u]$ defines the target range (the default setting is $[0.75, 0.9]$). This formulation provides maximum reward when occupancy falls within the optimal range, while creating a linear penalty gradient for both underutilization of space and overcrowding. When occupancy is zero, the reward equals -1, scaling linearly towards 1 as the occupancy approaches the lower threshold. Similarly, exceeding the upper threshold results in a linear decline from 1 toward -1 as occupancy approaches maximum capacity.

Reimbursement Reward

The reimbursement reward, R_{reimb} , introduces a distribution-aware approach by evaluating the facility's total daily revenue against statistical benchmarks derived from the referral dataset. This method captures both patient selection quality and volume:

$$R_{\text{reimb}}(R_{\text{total}}) = \begin{cases} \frac{R_{\text{total}}}{E[R]} - 1, & \text{if } R_{\text{total}} \leq E[R] \\ \frac{R_{\text{total}} - E[R]}{T[R] - E[R]}, & \text{if } R_{\text{total}} > E[R], \end{cases} \quad (5)$$

where R_{total} represents the facility's total daily reimbursement revenue, $E[R] = C_{\text{max}} \cdot \alpha \cdot \mu_r$ denotes the expected revenue benchmark (α occupancy at mean reimbursement rate μ_r), and $T[R] = C_{\text{max}} \cdot \beta \cdot Q_{90}(r)$ represents the target revenue benchmark (β occupancy at the 90th percentile reimbursement rate). Here, C_{max} is the maximum bed capacity, μ_r is the mean reimbursement rate from the referral distribution, and $Q_{90}(r)$ is the 90th percentile of the reimbursement rate distribution. In our implementation, we set $\alpha = 0.5$ and $\beta = T_l$ to represent realistic industry benchmarks. This formulation creates a continuous reward signal scaled to $[-1, 1]$ that accounts for the statistical properties of the referral pool, incentivizing optimal selection while maintaining appropriate census levels.

Nursing Cost Reward

The nursing cost reward, R_{nurse} , normalizes staffing costs relative to the theoretical maximum possible cost, with an additional bonus for maintaining optimal care hours per patient.

$$R_{\text{nurse}}(C_t, C_{\text{max}}, p_t) = \text{clip} \left(-2 \cdot \frac{C_t}{C_{\text{max}}} + 1 + B(p_t), -1, 1 \right) \quad (6)$$

where C_t is the current staffing cost, C_{max} is the maximum possible cost when all staff are scheduled, and $B(p_t)$ is an efficiency bonus that is defined as:

$$B(p_t) = \begin{cases} 0.2 \cdot \left(1 - \frac{|p_t - \bar{p}|}{\Delta p} \right) & \text{if } H_l \leq p_t \leq H_u \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

where p_t represents the mean nursing hours per patient at time t . The target nursing hours range $[H_l, H_u]$ defines the acceptable care-intensity range, with $\bar{p} = \frac{H_l + H_u}{2}$ and $\Delta p = \frac{H_u - H_l}{2}$.

The reward combines two components. The base term $-2 \cdot \frac{C_t}{C_{\text{max}}} + 1$ penalizes staffing expenditure linearly, ranging from 1 (at zero cost) to -1 (at maximum cost). The efficiency bonus $B(p_t)$ adds up to 0.2 when hours-per-patient fall within $[H_l, H_u]$. These terms encourage both cost efficiency and appropriate care intensity. The final value is clipped to $[-1, 1]$, and defaults to 0 when $C_{\text{max}} = 0$. It should be noted that $B(p_t)$ can be removed from the calculation of $R_{\text{nurse}}(C_t, C_{\text{max}}, p_t)$ with no other adjustments if the user does not wish to consider efficiency.

Rehospitalization Reward

The rehospitalization reward, $R_{\text{re Hosp}}$, is designed to penalize patient rehospitalizations while accounting for facility occupancy:

$$R_{\text{re Hosp}}(h, n, t) = \begin{cases} 1 & \text{if } h = 0 \\ 1 - 2 \cdot \left(\frac{h}{n} \right)^2 & \text{if } n > 0 \text{ and } h > 0 \\ -1 \cdot \min \left(\frac{t}{30}, 1 \right) & \text{if } n = 0 \end{cases} \quad (8)$$

where h is the number of rehospitalizations, n is the current number of patients, and t is the current time step. The reward function maximizes positive reinforcement when no patients are rehospitalized, creating a strong incentive to provide quality care and effectively manage the case mix. As the rate of rehospitalizations increases, the reward decreases quadratically, with the potential to reach a minimum value of -1 for extremely high rehospitalization rates. For facilities with zero patients, the penalty gradually increases over time, reaching a maximum negative reward of -1 after some number of time steps (defaulted to 30). This is put in place to allow an agent to amass patients at the beginning of new episodes. This approach balances the need to minimize rehospitalizations with maintaining an active patient population.

4 Applications and Baseline Results

In Section 4.1, we illustrate how SNFsim may be used to evaluate and compare hand-constructed policies in terms of their relative performance on different reward components. In Section 4.3, we evaluate and compare standard implementations of Proximal Policy Optimization (PPO) in SNFsim. We use a simplified rehospitalization classifier in this section, which is a lightweight alternative to the random forest model that preserves the directional relationships of the original model while guaranteeing monotonicity. While the random forest achieves high predictive fidelity (AUC of 0.921), its non-monotonic decision surface can introduce unpredictable risk signals during training, potentially impacting the agents ability to learn stable associations between actions and outcomes. The abstraction maintains the essential risk patterns and allows for controlled policy learning experiments. The simplified model considers LoS, mean CNA hours, reimbursement rate, and age, each contributing to the rehospitalization risk at varying degrees of influence consistent with the relationships identified in **Table 2**. Based on domain knowledge and the empirical findings of the regression analysis, we

posit that longer stays and higher CNA hours per patient are protective against rehospitalization, while higher reimbursement rates (a proxy for patient acuity) and advanced age are associated with elevated risk. The relative influence of each feature on rehospitalization risk is fixed in our simplified model but is fully configurable. Note that researchers employing SNFsim still have the option to use the original rehospitalization model discussed previously, or their own model, rather than the simplified version.

4. 1 Policy Analysis

Rolling out and comparing manually specified policies can provide insights into how different strategies balance or prioritize each of the competing objectives. This can help clarify the relationship among environment dynamics and primary objectives.

We ran five episodes for each policy with an episode length of 365 timesteps (equivalent to one year in the environment). Each facility had at their disposal 35 full-time CNAs, 10 PRN CNAs, and 15 agency CNAs. Each facility also had an optimal occupancy between 75% and 90%, was located in New York, received 10 referrals per day, and had 100 available beds.

We evaluate two sets of simple policies P_1 and P_2 and P_3 and P_4 . Policy P_1 focuses on optimizing occupancy and minimizing rehospitalizations by accepting every referral each day and keeping nursing hours per patient at as close to six hours as possible. Policy P_2 is more conservative, accepting just one referral per day (the one with the highest reimbursement rate), and hiring just one PRN and agency CNA per each daily shift. **Figure 8** displays the results from five rollouts of policies P_1 and P_2 within the `snf_v0` environment.

Policy P_3 emphasizes cost reduction (but not necessarily in an efficient manner) by accepting five random referrals each day and keeping nursing hours per patient at as close to a singular nursing hour per patient hours as possible. Policy P_4 prioritizes high-quality patient care by accepting the five highest-reimbursement (i.e., most intense) referrals each day. Additionally, P_4 maximizes CNA availability by increasing facility staffing and fully utilizing PRN and agency CNAs. **Figure 9** displays the results from five episodes of rolling out policies P_3 and P_4 within the `snf_v0` environment.

It is clear that by accepting every referral while maintaining a high level of nursing hours per patient, P_1 consistently achieves high occupancy as patients are admitted frequently and very few patients are rehospitalized due to the excellent levels of care provided. However, since the facility has 100 available beds and only ten referrals per day (all of which are accepted), there is no selectivity in acceptance. As a result, the reimbursement reward is lower than it could be if the facility received a larger number of referrals and selectively accepted those with higher reimbursement rates. Additionally, because P_1 ensures a high level of care by scheduling the CNAs required for high levels of patient care, the nursing cost management reward is also lower than its potential. Conversely, P_2 struggles to retain patients due to its low level of referral acceptance (just one per day), resulting in low occupancy and reimbursement rewards. Regardless, it manages to keep a low level of rehospitalizations despite spending very little on CNAs as there are fewer patients in the facility each day to care for.

Policy P_3 accepts five random referrals per day and allocates only a single nursing hour per patient, which is significantly below the level required for effective inpatient care. As a consequence, P_3 suffers from a high rate of rehospitalizations, low reimbursement (due to the frequent rehospitalizations), and consistently low occupancy. In contrast, policy P_4 selects the five highest-paying patients per day and maximizes CNA staffing. This approach results in substantial rewards in terms of reimbursement and rehospitalization prevention, however it is at the cost of a very low nursing cost management reward. Additionally, because the facility operates near full capacity due to the continual acceptance of referrals combined with high levels of care, the occupancy reward remains low due to the optimal occupancy being set to a value between 75% and 90%, as demonstrated in Equation 8.

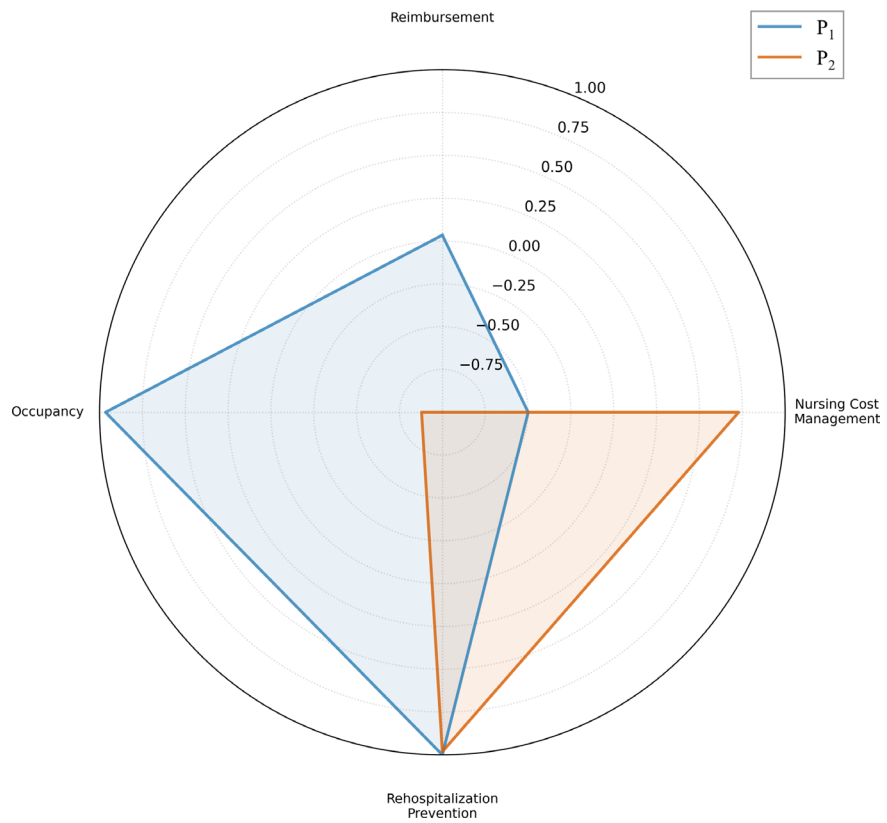


Figure 8. Performance comparison of five rollouts between P_1 and P_2 in the snf_v0 environment.

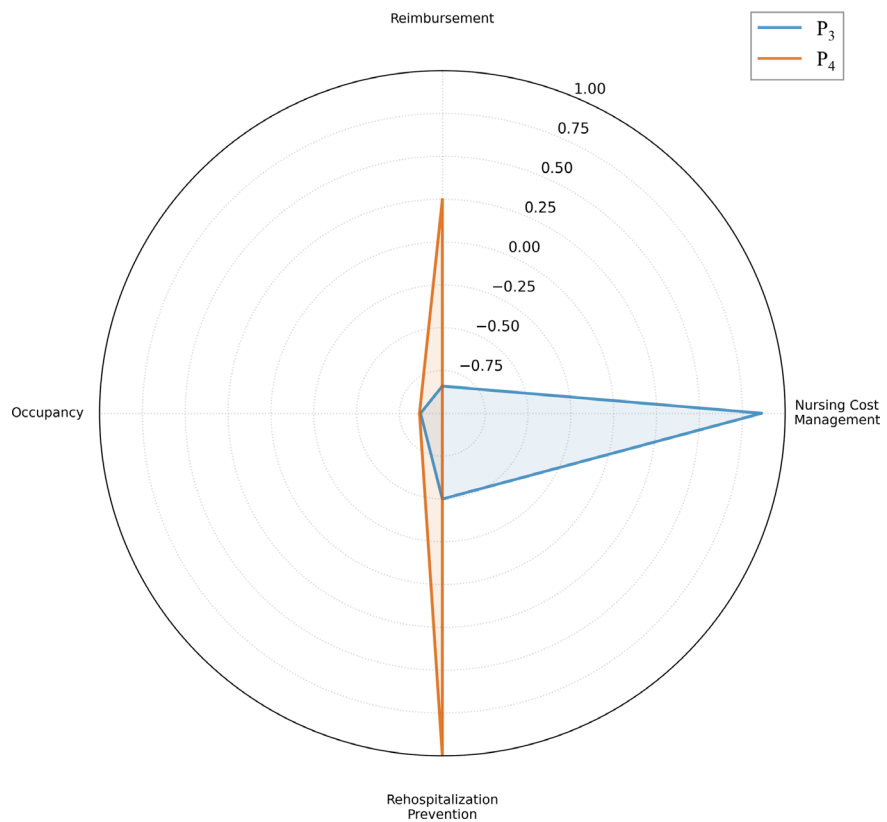


Figure 9. Performance comparison of five rollouts between P_3 and P_4 in the snf_v0 environment.

4. 2 Model Validation and Real-World Engagement

To build confidence in SNFsim, we conducted validation activities at multiple levels during its development. For patient-level and facility-level validation, notable measures included cross-validation of the rehospitalization risk model, calibration of staffing costs to Bureau of Labor statistics wage data, and validation that simulated patient demographics closely match empirical distributions from CMS claims data. System-level validation ensured that capacity constraints are enforced and boundary conditions are handled properly. As seen in Section 4.3, policies learned in a standard environment exhibit consistent behavior across multiple evaluation episodes, demonstrating stable facility dynamics.

To ensure the practical relevance of our framework, we have consulted with healthcare experts and frontline nursing staff through our industry partnership with *PointClickCare*, the leading provider of healthcare technology solutions for post-acute care. Our goal was to obtain an understanding of their decision-making priorities. These discussions revealed that facility operators generally face three primary challenges: 1) balancing quality metrics mandated by CMS with financial sustainability, 2) making rapid staffing decisions despite labor shortages, and 3) handling the tradeoff of the acceptance of high-acuity patients and maintaining low rehospitalization rates to ensure a continued positive reputation and quality care for patients.

Both the simulator itself and the weight configurations examined in Section 4.3 were designed to reflect these real-world priorities. However, it should be noted that in its current state, this work represents a computational framework rather than a deployment-ready tool.

Moving from this testbed to practical adoption will require addressing several unique implementation challenges. First, building trust in recommendations driven by machine learning requires transparent explainability tools that show why specific referrals are recommended for acceptance or rejection or why certain staffing decisions are suggested. Next, validation through pilot deployments (in which recommendations are made alongside human decision-makers without influencing operations) will be non-negotiable in ensuring user confidence. Finally, to maximize usefulness, individual facilities should calibrate the simulator to best reflect their respective transition probabilities. As SNFsim is highly modular, this would require large-scale historical data and increased operational granularity built into simulation.

4. 3 RL Methods

In this section, we train multiple RL agents within SNFsim and examine how the resulting policies differ in terms of decision-making and objective balancing.

We evaluate the effectiveness of PPO within SNFsim due to its stability in training under noisy reward signals and effectiveness in high-dimensional state spaces like ours. PPO aims to maximize a cumulative reward signal by learning a policy function while ensuring stable learning through constrained policy updates (*Schulman et al., 2017*). The algorithm maximizes the clipped surrogate objective:

$$L^{CLIP}(\theta) = \hat{\mathbb{E}}_t \left[\min \left(r_t(\theta) \hat{A}_t, \text{clip}(r_t(\theta), 1 - \epsilon, 1 + \epsilon) \hat{A}_t \right) \right], \quad (9)$$

where $r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)}$ is the probability ratio between the new and old policies (π_θ and $\pi_{\theta_{old}}$ respectively), \hat{A}_t is the estimated advantage function, and ϵ is a hyperparameter that constrains policy updates.

For simplicity within our multi-objective problem, we employ a fixed-weight scalarization approach that converts the vector of rewards into a single scalar value through weighted summation. With m objectives (where $m = 4$ in our case), the scalar reward at each timestep is calculated as:

$$R(s, a) = \sum_{i=1}^m w_i R_i(s, a) \quad (10)$$

where w_i represents the weight assigned to the i^{th} objective and $R_i(s, a) \in [-1, 1]$ is the normalized reward for that objective. This approach allows us to prioritize different objectives by adjusting their respective weights.

4. 3. 1 Method Performance

We evaluated three PPO agents (hyperparameters in **Table A14**) with distinct reward weight configurations to explore different policy priorities within SNFsim. **Table 4** summarizes the key performance metrics and **Figure 10** shows training performance under varied weight configurations.

The care-focused agent achieved the highest absolute performance, with final-quarter rewards approaching 75% of the theoretical maximum (2.94 out of 4.0). This agent also exhibited the steepest learning curve and attained the highest average reward (2.54).

In contrast, the financially-focused agent (**Figure 10c**) delivered lower absolute rewards but demonstrated the most dramatic relative improvement (312.67%). This indicates that financial optimization requires more sophisticated strategies. Despite this improvement, the agent reached only approximately 13% of the theoretical maximum reward of 4, which hints at difficulty in simultaneously optimizing financial components without considering the importance of patient wellness and facility occupancy.

The balanced agent (**Figure 10a**) achieved moderate performance with a 42.74% improvement rate and a trend slope of approximately 0.60, reflecting the trade-offs between competing objectives. These results suggest that policy optimization in multi-dimensional SNF management should consider the difficulty of improving different objectives and prioritize accordingly based on clear individual goals.

Table 5 shows clear tradeoffs between policies trained with different objective weightings. The care-focused policy achieved superior performance across both quality and overall financial metrics: 13.6% rehospitalization rate (approximately half the real-world SNF average of 23.5% (**Minges et al., 2019**), 73.0% occupancy (the highest), and \$28,534 daily profit (16% higher than Balanced, 13% higher than Financial Focus). This policy invested in staffing (\$9,933 daily cost, 47% above other policies) and selectively accepted referrals (60.1% acceptance rate), allowing it to maintain higher occupancy with patients that better fit the care potential of the SNF.

Conversely, the financial-focused policy learned to accept all referrals (100% acceptance rate) and minimized staffing costs (\$6,922), but suffered from the highest rehospitalization rate (42.8%) and, as a result, the lowest occupancy (53.9%). The high patient turnover from inadequate care prevented sustained growth, ultimately resulting in lower profit despite lower costs.

The Balanced policy achieved intermediate outcomes (37.3% rehospitalization; approximately 1.6 times the national average, 80% acceptance rate, \$24,680 profit), but was dominated by the care-focused approach on both quality and financial dimensions.

These results suggest that prioritizing quality through adequate staffing and selective admissions generates superior financial outcomes compared to cost minimization, as preventing rehospitalizations results in more overall value than reducing care expenses.

Table 4. Summary of training performance across different PPO weight configurations.

Metric	Balanced	Financial Focus	Care Focus
Weights	[1.0, 1.0, 1.0, 1.0]	[1.6, 1.6, 0.4, 0.4]	[0.4, 0.4, 1.6, 1.6]
Average Reward	1.41	0.40	2.54
First Quarter Avg	1.07	0.13	1.88
Last Quarter Avg	1.53	0.53	2.94
Improvement (%)	42.74%	312.67%	56.00%
Trend Slope	0.5982	0.5307	1.3929

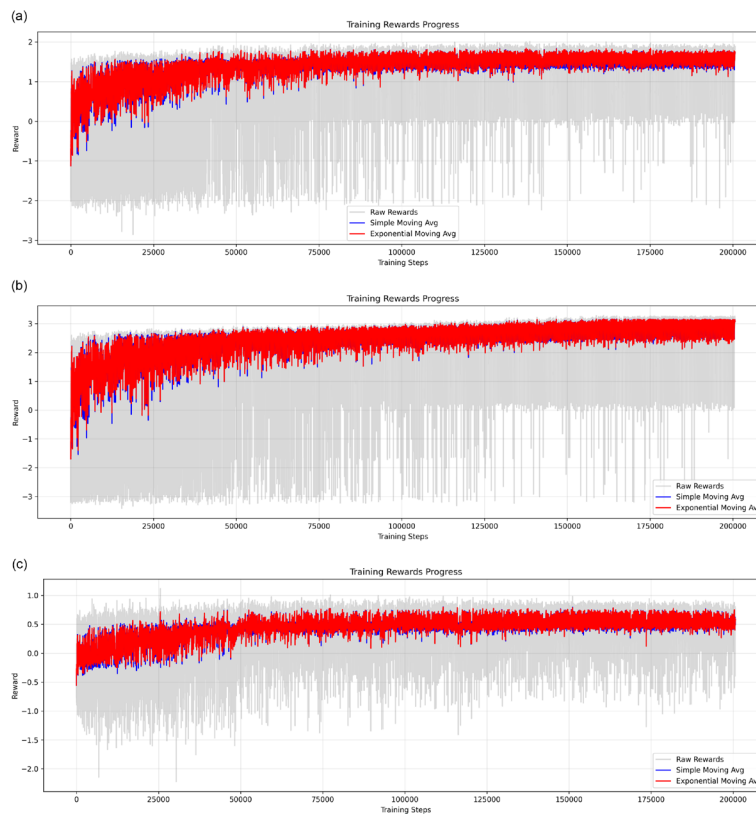


Figure 10. PPO agent training performance under varied weight configurations across 200,000 timesteps. Objective weights follow the order [reimbursement, nursing costs, rehospitalization, occupancy]. (a) Balanced approach with uniform weights [1.0, 1.0, 1.0, 1.0]. (b) Care optimization with weights [0.4, 0.4, 1.6, 1.6], emphasizing rehospitalization minimization and occupancy management while de-emphasizing financial metrics. (c) Financial optimization with weights [1.6, 1.6, 0.4, 0.4], prioritizing reimbursement maximization and nursing cost minimization over patient care and facility occupancy. The theoretical summed reward range is [-4, 4].

Table 5. Outcomes by policy configuration over 10 different 365-day episodes. Weight vectors indicate a priori selection of importance for [reimbursement, nursing cost, rehospitalization, occupancy] objectives.

Metric	Balanced	Financial Focus	Care Focus
Weights	[1.0, 1.0, 1.0, 1.0]	[1.6, 1.6, 0.4, 0.4]	[0.4, 0.4, 1.6, 1.6]
Occupancy Rate	55.62% ± 0.24%	53.89% ± 0.20%	72.97% ± 0.39%
Rehospitalization Rate	37.3% ± 4.1%	42.8% ± 3.6%	13.6% ± 3.4%
Referral Acceptance Rate	80.0%	100.0%	60.1%
Daily Revenue	\$31,618 ± \$131	\$32,199 ± \$136	38,467 ± \$141
Daily Cost	\$6,938 ± \$1.81	\$6,922 ± \$1.75	\$9,933 ± \$1.44
Daily Profit	\$24,680	\$25,277	\$28,534

5 Discussion

SNFsim provides a flexible open-source platform calibrated using real-world datasets for the development and testing of complex RL algorithms and policies that tackle long-horizon, multi-dimensional, and multi-objective optimization challenges. By simplifying many of the low-level details, SNFsim focuses instead on the high-level interactions occurring within SNFs, allowing researchers and healthcare professionals to explore the potential outcomes of various policy implementations and

management strategies in a controlled, risk-free environment. By simulating real-world conditions and constraints, the platform provides a sandbox for identifying and addressing the limitations of current algorithms, encouraging the advancement of RL strategies that are applicable in real-world SNF settings.

Amidst the complexities of complex stochastic real-world environments in which algorithms face the risk of non-convergence or settling to local optima, SNFsim provides a valuable framework for constructing custom environments rooted in real-world healthcare data. We demonstrate an example of this process in Section 3.6, in which we build a custom Gymnasium environment, `snf_v0`, on top of SNFsim and later train RL agents within that custom environment.

It is important to note that although SNFsim incorporates predictive components (such as the rehospitalization risk model), it is not intended to function as a validated forecasting tool, nor do we claim complete empirical predictive accuracy in this work. Rather, the simulator provides a controlled environment in which the consequences of different staffing and referral policies can be explored through forward simulation. As such, SNFsim enables users to explore how alternative decisions may influence revenue, occupancy, staffing costs, and rehospitalizations under consistent assumptions. Suggestions on how SNFsim or a similar system might be implemented to aid in real-world decision-making is discussed in Section 4.2.

5. 1 Relevance to RL Research

RL applications are commonly developed within single-objective environments, focusing on optimizing a scalar reward signal. However, healthcare inherently involves balancing multiple objectives, which can be a significant challenge. Discrete-event simulators have long been a tool for multi-objective optimization in healthcare (*Wang et al., 2015; Al-Hawari et al., 2022*), providing a framework for decision-making within complex systems with competing priorities. By taking into account the inter-related outcomes of occupancy, net revenue, and rehospitalization, and modeling their nonlinear relationships, SNFsim allows for the development of more robust and adaptable multi-objective RL solutions.

Additionally, the infinite-horizon of decision-making in SNFs distinguishes it from many conventional RL applications, which are often based on finite horizons with clearly defined start and endpoints based on pre-defined goals (e.g., reaching the end of a maze or landing a helicopter directly in the center of a helipad). In SNFs, there is no fixed endpoint; the facility must continuously operate and evolve, making it imperative for RL policies to focus on long-term sustainability and adaptability. Using simulators such as SNFsim for infinite-horizon problems offers significant benefits. Simulators provide an interactive environment where algorithms can be continuously trained, tested, and refined under an assortment of conditions that may not be fully represented or available in offline datasets. This is particularly beneficial for infinite-horizon problems, where the decision-making process extends indefinitely into the future, and the system must adapt to evolving states and objectives over time. SNFsim allows for the exploration of long-term strategies and the evaluation of their impacts in a controlled setting, enabling the development of more robust and effective RL solutions that are better equipped to handle the complexities and uncertainties of real-world applications.

Finally, the interdisciplinary nature of RL in SNFs, involving collaboration across healthcare, operations management, and AI, not only enhances the applicability of RL solutions, but fosters a broader understanding of real-world complexities and allows for the exploration of more pragmatic RL applications.

The exploration of RL within the context of SNFs is not just a step towards improving operational efficiency and patient care in SNFs; it is a significant contribution to the field of RL as a whole. It challenges existing ideas, introduces new problems, and could help bridge the gap between methodological research and practical applications.

5. 2 Relevance to healthcare Decision-making

The use of SNFsim in conjunction with RL has the potential to enhance operational efficiency and resource management. RL models, trained and tested within a simulated SNF environment that can be tailored to an individual SNF, can analyze complex patterns and predict future trends. This predictive capability could support more informed decision-making regarding critical aspects like staffing levels,

resource allocation, and patient admissions, and thereby help to achieve operational efficiency and high-quality patient care.

Our simulator provides a safe and controlled environment for testing and refining RL-based strategies. In the sensitive realm of healthcare, direct experimentation can be risky and ethically problematic. A simulated environment using de-identified healthcare data allows for extensive testing of different policies and decision-making scenarios without risk to patients or facilities. This feature is invaluable for validating RL models and ensuring their reliability and safety before real-world implementation.

Finally, the predictive analytics capability of a discrete-event simulator like SNFsim is particularly relevant for proactive healthcare management. By identifying potential risks and anticipating future scenarios, RL models can advise on preventive measures and strategic adjustments. For instance, predicting high-risk patients for rehospitalization and planning appropriate interventions can significantly improve patient outcomes and reduce the financial and reputational risk associated with hospital rehospitalizations.

6 Future Work

The goal of SNFsim and the custom `snf_v0` environment is to provide a testbed for the future development of RL methodology and decision support tools, which we discuss below; however, there are avenues for future research and enhancements of the simulator and environment themselves.

In the future, we aim to expand upon the multi-objective nature of `snf_v0` by incorporating additional objectives that reflect the competing goals inherent in healthcare facility management. For instance, we plan to incorporate outcomes such as patient satisfaction and staff well-being, providing a more comprehensive approach to decision-making. Additionally, adapting the environment to dynamically adjust these objectives based on changing circumstances or policies could offer a more realistic simulation of healthcare operations.

Future iterations could also explore more nuanced representations of the state and action spaces, capturing additional complexities of real-world SNFs. This could include more detailed patient profiles and a broader range of operational decisions. Simulating a multi-agent ecosystem, representing a group of facilities such as those in a county, or state, could also prove advantageous. In this setup, each agent would manage the decision-making for one facility. These facilities might each have distinct priorities and should learn to make strategic choices based on their understanding of the likely actions of other facilities within their ecosystem.

In the current implementation, SNFsim models referral arrivals as a Poisson process with constant rate parameter, which is effective for facilities operating in stable markets with established hospital referral relationships, though referral patterns in the real-world often vary over time based on a variety of facility-level features as well as hospital partnerships (McHugh et al., 2021; Kim et al., 2019). As such, future work should consider the task of capturing shifting referral demographics and numbers over time, ensuring to capture the relationship between geographical location, facility-level features, and hospital relationship on sampled referrals.

While SNFsim uses real-world data collected from SNFs, access to additional anonymized data could improve the realism and applicability of the simulator by enabling more accurate modeling and policy development. For instance, gaining access to staff data could enable us to integrate employee experience into the simulator, thereby facilitating more refined and informed staffing decisions. Similarly, acquiring comprehensive facility surveys filled out by patients or their relatives would facilitate the integration of a scoring system within the simulator. This addition could allow actions to impact a perceived facility score, potentially influencing both the quality and volume of generated referrals.

Ultimately, SNFsim is a valuable resource in healthcare decision-making, providing a modular platform calibrated using real-world medical data and offering researchers and practitioners an opportunity to explore and develop techniques for effectively balancing multiple objectives in SNFs.

Ethics

The study, along with a waiver of consent for secondary analysis of deidentified data, was approved by the Western University Health Sciences Research Ethics Board under project ID 114989. The Western University Health Sciences Research Ethics Board operates within the provisions of the Ontario Personal Health Information Protection Act (2004) and its applicable regulations and is registered with

the US Department of Health & Human Services under institutional review board registration number institutional review board 00000940.

ORCID iDs

Caroline Strickland  <https://orcid.org/0000-0003-2458-3848>

Brittin Wagner  <https://orcid.org/0000-0002-3518-4546>

Stanley Wang  <https://orcid.org/0009-0008-4681-5420>

Daniel J Lizotte  <https://orcid.org/0000-0002-9258-8619>

Funding

This work was supported in part by funding from the Natural Sciences and Engineering Research Council of Canada and from PointClickCare. (NSERC Alliance grant ALLRP 566302-21).

Conflict of Interest

No competing interests reported.

Data and code availability

The raw data sets analyzed during this study are not publicly available as they are proprietary. However, synthetic data generated in this paper is included in the base repository. The base SNFsim code is available at <https://github.com/Scorks/SNFsim>.

References

- Acumen.** 2018. Skilled nursing facilities patient-driven payment model technical report. http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/SNFPPS/Downloads/PDPM_Technical_Report_508.pdf [Accessed August 29, 2024].
- Ahangar RR.** 2022. Computation, modeling, and simulation of HIV-AIDS epidemics with vaccination. *Journal of Applied Mathematics and Physics* **10**: 1066–1082. DOI: <https://doi.org/10.4236/jamp.2022.104073>
- Al-Hawari T, Alrejjal A, Mumani AA, Momani A, Alhawari H.** 2022. A framework for multi-response optimization of healthcare systems using discrete event simulation and response surface methodology. *Arabian Journal for Science and Engineering* **47**: 15001–15014. DOI: <https://doi.org/10.1007/s13369-022-06633-8>
- Butler TW, Karwan KR, Sweigart JR, Reeves GR.** 1992. An integrative model-based approach to hospital layout. *IIE Transactions* **24**: 144–152. DOI: <https://doi.org/10.1080/07408179208964211>
- Centers for Medicare and Medicaid Services.** Patient driven payment model. <https://www.cms.gov/medicare/medicare-fee-for-service-payment/snfpps/pdpm.ca> [Accessed June 10, 2023].
- Chandra A, Rahman PA, Sneve A, McCoy RG, Thorsteinsdottir B, Chaudhry R, Storlie CB, Murphree DH Jr, Hanson GJ, Takahashi PY.** 2019. Risk of 30-day hospital readmission among patients discharged to skilled nursing facilities: Development and validation of a risk-prediction model. *Journal of the American Medical Directors Association* **20**: 444–450. DOI: <https://doi.org/10.1016/j.jamda.2019.01.137>
- England W, Roberts SD.** 1978. Applications of computer simulation to health care. *Technical report, IEEE.*
- Günal MM, Pidd M.** 2010. Discrete event simulation for performance modelling in health care: a review of the literature. *Journal of Simulation* **4**: 42–51. DOI: <https://doi.org/10.1057/jos.2009.25>
- Hambly B, Xu R, Yang H.** 2023. Recent advances in reinforcement learning in finance. *Mathematical Finance* **33**: 437–503. DOI: <https://doi.org/10.1111/mafi.12382>
- Jalalimanesh A, Shahabi Haghghi H, Ahmadi A, Soltani M.** 2017. Simulation-based optimization of radiotherapy: Agent-based modeling and reinforcement learning. *Mathematics and Computers in Simulation* **133**: 235–248. DOI: <https://doi.org/10.1016/j.matcom.2016.05.008>
- Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Möller J.** 2012. Modeling using discrete event simulation: a report of the ISPOR-SMDM modeling good research practices task force-4. *Medical Decision Making* **32**: 701–711. DOI: <https://doi.org/10.1177/0272989X12455462>
- Kim KL, Li L, Kuang M, Horwitz LI, Desai SM.** 2019. Changes in hospital referral patterns to skilled nursing facilities under the hospital readmissions reduction program. *Medical Care* **57**: 695–701. DOI: <https://doi.org/10.1097/MLR.0000000000001169>
- Kolesar P.** 1970. A markovian model for hospital admission scheduling. *Management Science* **16**: B384. DOI: <https://doi.org/10.1287/mnsc.16.6.B384>
- Laes-Kushner R.** 2018. Skilled nursing facilities: Too many beds. <https://repository.escholarship.umassmed.edu/handle/20.500.14038/26962> [Accessed September 25, 2024].
- Lou Z, Hass Z, Kong N.** 2025. An interpretable machine learning study for developing a binary classifier for predicting rehospitalization from skilled nursing facilities. *Healthcare Analytics* **7**: 100387. DOI: <https://doi.org/10.1016/j.health.2025.100387>

- Man CD**, Micheletto F, Lv D, Breton M, Kovatchev B, Cobelli C. 2014. The UVA/PADOVA Type 1 diabetes simulator: new features. *Journal of Diabetes Science and Technology* **8**: 26–34. DOI: <https://doi.org/10.1177/1932296813514502>
- Marshall DA**, Burgos-Liz L, IJerman MJ, Crown W, Padula WV, Wong PK, Pasupathy KS, Higashi MK, Osgood ND, ISPOR Emerging Good Practices Task Force. 2015. Selecting a dynamic simulation modeling method for health care delivery research-part 2: report of the ISPOR dynamic simulation modeling emerging good practices task force. *Value in Health* **18**: 147–160. DOI: <https://doi.org/10.1016/j.jval.2015.01.006>
- McHugh JP**, Rapp T, Mor V, Rahman M. 2021. Higher hospital referral concentration associated with lower-risk patients in skilled nursing facilities. *Health Services Research* **56**: 839–846. DOI: <https://doi.org/10.1111/1475-6773.13654>
- Minges KE**, Campbell Britton M, Clark BW, Ouellet GM, Hodshon B, Chaudhry SI. 2019. Hospital readmission from skilled nursing facilities (SNFs): perspectives of hospital and SNF providers. *Journal of the American Medical Directors Association* **20**: 1050–1051. DOI: <https://doi.org/10.1016/j.jamda.2019.03.005>
- Mnih V**, Kavukcuoglu K, Silver D, Graves A, Antonoglou I, Wierstra D, Riedmiller M. 2013. Playing atari with deep reinforcement learning. DOI: <https://doi.org/ArXiv:1312.5602>
- Mor V**, Intrator O, Feng Z, Grabowski DC. 2010. The revolving door of rehospitalization from skilled nursing facilities. *Health Affairs* **29**: 57–64. DOI: <https://doi.org/10.1377/hlthaff.2009.0629>
- Nagendra S**, Podila N, Ugarakhod R, George K. 2017. Comparison of reinforcement learning algorithms applied to the cart-pole problem. 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI). IEEE. DOI: <https://doi.org/10.1109/ICACCI.2017.8125811>
- Nursa**. 2024. Cna salary data by state. <https://nursa.com/salary> [Accessed March 1, 2024].
- Ozcan YA**, Wogen SE, Mau LW. 1998. Efficiency evaluation of skilled nursing facilities. *Journal of Medical Systems* **22**: 211–224: 211224. DOI: <https://doi.org/10.1023/a:1022657600192>
- Patki N**, Wedge R, Veeramachaneni K. 2016. The Synthetic Data Vault. 2016 IEEE International Conference on Data Science and Advanced Analytics. Montreal, QC, Canada. DOI: <https://doi.org/10.1109/DSAA.2016.49>
- Pauly V**, Mendizabal H, Gentile S, Auquier P, Boyer L. 2019. Predictive risk score for unplanned 30-day rehospitalizations in the French universal health care system based on a medico-administrative database. *PLoS One* **14**: e0210714. DOI: <https://doi.org/10.1371/journal.pone.0210714>
- Petersen BK**, Yang J, Grathwohl WS, Cockrell C, Santiago C, An G, Faissol DM. 2018. Precision medicine as a control problem: Using simulation and deep reinforcement learning to discover adaptive, personalized multi-cytokine therapy for sepsis. DOI: <https://doi.org/ArXiv:1802.10440>
- Pidd M**. 1997. Tools for thinking: Modelling in management science. *Journal of the Operational Research Society* **48**. DOI: <https://doi.org/10.2307/3010517>
- Rahman M**, Grabowski DC, Mor V, Norton EC. 2016. Is a skilled nursing facility's rehospitalization rate a valid quality measure? *Health Services Research* **51**: 2158–2175. DOI: <https://doi.org/10.1111/1475-6773.12603>
- Rising EJ**, Baron R, Averill B. 1973. A systems analysis of a university-health-service outpatient clinic. *Operations Research* **21**: 1030–1047. DOI: <https://doi.org/10.1287/opre.21.5.1030>
- Robinson GH**, Wing P, Davis LE. 1968. Computer simulation of hospital patient scheduling systems. *Health Services Research* **3**: 130–141.
- Rutter CM**, Ozik J, DeYoreo M, Collier N. 2019. Microsimulation model calibration using incremental mixture approximate bayesian computation. *The Annals of Applied Statistics* **13**: 2189–2212. DOI: <https://doi.org/10.1214/19-aos1279>
- Schulman J**, Wolski F, Dhariwal P, Radford A, Klimov O. 2017. Proximal policy optimization algorithms. arXiv.
- Silver D**, Hubert T, Schrittwieser J, Antonoglou I. 2017. Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm. arXiv.
- Sklar M**. 1959. Fonctions de rpartition n dimensions et leurs marges. *Annales de IISUP*. Vol. 8 : 229231.
- Spielauer M**. 2011. What is social science microsimulation? *Social Science Computer Review* **29**: 9–20. DOI: <https://doi.org/10.1177/0894439310370085>
- Stafylopatis A**, Blekas K. 1998. Autonomous vehicle navigation using evolutionary reinforcement learning. *European Journal of Operational Research* **108**: 306–318. DOI: [https://doi.org/10.1016/S0377-2217\(97\)00372-X](https://doi.org/10.1016/S0377-2217(97)00372-X)
- Strickland C**, Chi N, Ditz L, Gomez L, Wagner B, Wang S, Lizotte DJ. 2023. Factors influencing admission decisions in skilled nursing facilities: retrospective quantitative study. *Journal of Medical Internet Research* **25**: e43518. DOI: <https://doi.org/10.2196/43518>
- Sun X**, Kong N, Zhao Y, Sakib N, Meng C, Meng H, Hyer K, Li Y, Masterson C, Li M. 2023. A latent survival model integrated computer simulation-based evaluation for nursing home staffing. *Computers & Industrial Engineering* **177**: 109074. DOI: <https://doi.org/10.1016/j.cie.2023.109074>
- Tracy M**, Cerdá M, Keyes KM. 2018. Agent-based modeling in public health: Current applications and future directions. *Annual Review of Public Health* **39**: 77–94. DOI: <https://doi.org/10.1146/annurev-publhealth-040617-014317>
- Wang Y**, Hay Lee L, Peng Chew EK, Shao S. 2015. Multi-objective optimization for a hospital inpatient flow process via Discrete Event Simulation. 2015 Winter Simulation Conference. Huntington Beach, CA, USA. DOI: <https://doi.org/10.1109/WSC.2015.7408521>
- Williamson K**. 2024. Skilled nursing facility (snf) billing guidelines 2024. <https://tranquilmedsolutions.com/skilled-nursing-facility-billing-services> [Accessed July 7, 2026].

Appendices

Table A1. PDPM payment groups to code value.

PT/OT	SLP	NURS	NPG	Code Value
TA	SA	ES3	NA	A
TB	SB	ES2	NB	B
TC	SC	ES1	NC	C
TD	SD	HDE2	ND	D
TE	SE	HDE1	NE	E
TF	SF	HBC2	NF	F
TG	SG	CBC2		G
TH	SH	CA2		H
TI	SI	CBC1		I
TJ	SJ	CA1		J
TK	SK	BAB2		K
TL	SL	BAB1		L
TM		HBC1		M
TN		LDE2		N
TO		LDE1		O
TP		LBC2		P
		LBC1		Q
		CDE2		R
		CDE1		S
		PDE2		T
		PDE1		U
		PBC2		V
		PA2		W
		PBC1		X
		PA1		Y

Table A2. PT and OT case mix groups and PT and OT CMI based on clinical category and PT and OT function score.

Clinical Category	PT & OT Function Score	PT & OT Case Mix Group	PT CMI	OT CMI
Major Joint Replacement or Spinal Surgery	0-5	TA	1.53	1.49
Major Joint Replacement or Spinal Surgery	6-9	TB	1.69	1.63
Major Joint Replacement or Spinal Surgery	10-23	TC	1.88	1.68
Major Joint Replacement or Spinal Surgery	24	TD	1.92	1.53
Other Orthopedic	0-5	TE	1.42	1.41
Other Orthopedic	6-9	TF	1.61	1.59

Clinical Category	PT & OT Function Score	PT & OT Case Mix Group	PT CMI	OT CMI
Other Orthopedic	10-23	TG	1.67	1.64
Other Orthopedic	24	TH	1.16	1.15
Medical Management	0-5	TI	1.13	1.17
Medical Management	6-9	TJ	1.42	1.44
Medical Management	10-23	TK	1.52	1.54
Medical Management	24	TL	1.09	1.11
Non-Orthopedic Surgery and Acute Neurologic	0-5	TM	1.27	1.30
Non-Orthopedic Surgery and Acute Neurologic	6-9	TN	1.48	1.49
Non-Orthopedic Surgery and Acute Neurologic	10-23	TO	1.55	1.55
Non-Orthopedic Surgery and Acute Neurologic	24	TP	1.08	1.09

Table A3. PDPM assessment type to code value.

Assessment Type	Code Value
Initial Patient Assessment	0
PPS 5-Day Assessment	1

Table A4. PT and OT case mix groups and PT and OT CMIs based on clinical category and PT and OT function score.

Clinical Category	PT & OT Function Score	PT & OT Case Mix Group	PT CMI	OT CMI
Major Joint Replacement or Spinal Surgery	0-5	TA	1.53	1.49
Major Joint Replacement or Spinal Surgery	6-9	TB	1.69	1.63
Major Joint Replacement or Spinal Surgery	10-23	TC	1.88	1.68
Major Joint Replacement or Spinal Surgery	24	TD	1.92	1.53
Other Orthopedic	0-5	TE	1.42	1.41
Other Orthopedic	6-9	TF	1.61	1.59
Other Orthopedic	10-23	TG	1.67	1.64
Other Orthopedic	24	TH	1.16	1.15
Medical Management	0-5	TI	1.13	1.17
Medical Management	6-9	TJ	1.42	1.44
Medical Management	10-23	TK	1.52	1.54
Medical Management	24	TL	1.09	1.11
Non-Orthopedic Surgery and Acute Neurologic	0-5	TM	1.27	1.30
Non-Orthopedic Surgery and Acute Neurologic	6-9	TN	1.48	1.49
Non-Orthopedic Surgery and Acute Neurologic	10-23	TO	1.55	1.55

Clinical Category	PT & OT Function Score	PT & OT Case Mix Group	PT CMI	OT CMI
Non-Orthopedic Surgery and Acute Neurologic	24	TP	1.08	1.09

Table A5. SLP case mix groups and SLP CMIs based on whether patient has a mechanically altered diet or swallowing disorder and the presence of acute neurological conditions, SLP-related comorbidity, or cognitive impairment.

Condition*	Mechanically Altered Diet or Swallowing Disorder	SLP Case Mix Group	SLP CMI
None	Neither	SA	0.68
None	Either	SB	1.82
None	Both	SC	2.66
Any One	Neither	SD	1.46
Any One	Either	SE	2.33
Any One	Both	SF	2.97
Any Two	Neither	SG	2.04
Any Two	Either	SH	2.85
Any Two	Both	SI	3.51
All Three	Neither	SJ	2.98
All Three	Either	SK	3.69
All Three	Both	SL	4.19

*Presence of Acute Neurological Condition, SLP-Related Comorbidity, or Cognitive Impairment.

Table A6. Nursing payment group (CMG) and corresponding CMIs based on RUG-IV Nursing RUG, extensive services status, clinical conditions, depression status, and restorative nursing services.

RUG-IV Nursing RUG	Extensive Services	Clinical Conditions	Depression	RNS	Function Score	CMG	CMI
ES3	Trach & Ventilator				0-14	ES3	4.04
ES2	Trach or Ventilator				0-14	ES2	3.06
ES1	Infection Isolation				0-14	ES1	2.91
HE2/HD2		SMC	Yes		0-5	HDE2	2.39
HE1/HD1		SMC	No		0-5	HDE1	1.99
HC2/HB2		SMC	Yes		6-14	HBC2	2.23
HC1/HB1		SMC	No		6-14	HBC1	1.85
LE2/LD2		RMC	Yes		0-5	LDE2	2.07
LE1/LD1		RMC	No		0-5	LDE1	1.72
LC2/LB2		RMC	Yes		6-14	LBC2	1.71
LC1/LB1		RMC	No		6-14	LBC1	1.43
CE2/CD2		CRC	Yes		0-5	CDE2	1.86
CE1/CD1		CRC	No		0-5	CDE1	1.62
CC2/CB2		CRC	Yes		6-14	CBC2	1.54
CA2		CRC	Yes		15-16	CA2	1.08
CC1/CB1		CRC	No		6-14	CBC1	1.34
CA1		CRC	No		15-16	CA1	0.94

RUG-IV Nursing RUG	Extensive Services	Clinical Conditions	Depression	RNS	Function Score	CMG	CMI
BB2/BA2		BCS		2+	11-16	BAB2	1.04
BB1/BA1		BCS		0-1	11-16	BAB1	0.99
PE2/PD2		ADL		2+	0-5	PDE2	1.57
PE1/PD1		ADL		0-1	0-5	PDE1	1.47
PC2/PB2		ADL		2+	6-14	PBC2	1.21
PA2		ADL		2+	15-16	PA2	0.7
PC1/PB1		ADL		0-1	6-14	PBC1	1.13
PA1		ADL		0-1	15-16	PA1	0.66

Table A7. Non-Therapy Ancillaries (NTA) CMIs based on NTA case mix groups and NTA score ranges.

NTA Score Range	NTA Case Mix Group	CMI
12+	NA	3.25
9-11	NB	2.53
6-8	NC	1.85
3-5	ND	1.34
1-2	NE	0.96
0	NF	0.72

Table A8. Urban rate components.

Rate Component	PT	OT	SLP	Nursing	NTA	Non-Case-Mix (NCM)
Per Diem Amount	\$62.84	\$58.49	\$23.46	\$109.55	\$82.64	\$98.10

Table A9. Rural rate components.

Rate Component	PT	OT	SLP	Nursing	NTA	Non-Case-Mix (NCM)
Per Diem Amount	\$71.63	\$65.79	\$29.56	\$104.66	\$78.96	\$99.91

Table A10. Day in stay adjustment factor.

Day in Stay	Adjustment Factor
1-20	1.00
21-27	0.98
28-34	0.96
35-41	0.94
42-48	0.92
49-55	0.90
56-62	0.88
63-69	0.86
70-76	0.84
77-83	0.82
84-90	0.80
91-97	0.78
98-150	0.76

Table A11. NTA component adjustment factor.

Day in Stay	Adjustment Factor
1-3	3.00
4-150	1.00

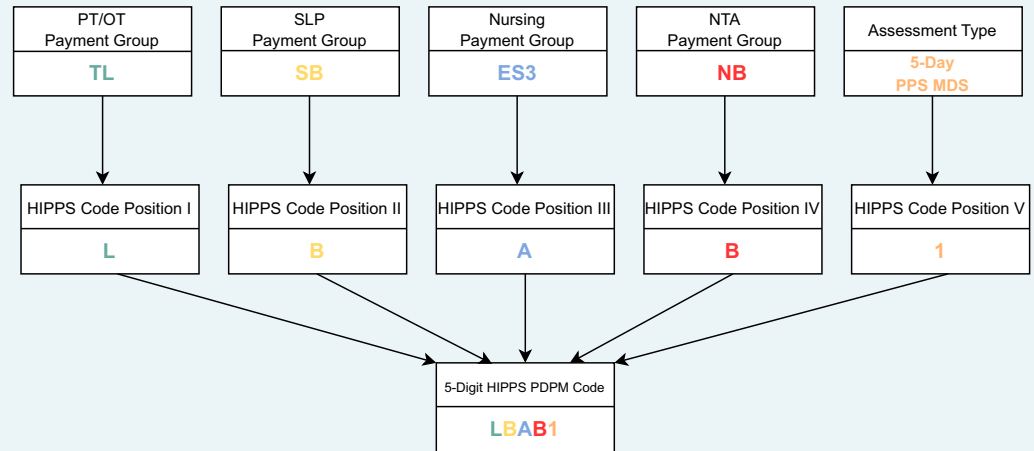


Figure A1. PDPM HIPPS code classification build sample.

Table A12. Estimated Hourly Wages for Full-time, PRN, and Agency CNAs by State based on (Nursa, 2024).

State	Full-time CNA (\$/hr)	PRN CNA (\$/hr)	Agency CNA (\$/hr)
Alabama	15.04	18.05	20.30
Alaska	22.63	27.16	30.55
Arizona	19.69	23.63	26.58
Arkansas	15.41	18.49	20.80
California	22.63	27.16	30.55
Colorado	20.95	25.14	28.28
Connecticut	20.70	24.84	27.95
Delaware	18.57	22.28	25.07
Florida	17.67	21.20	23.85
Georgia	16.77	20.12	22.64
Hawaii	21.63	25.96	29.20
Idaho	17.92	21.50	24.19
Illinois	19.87	23.84	26.82
Indiana	18.10	21.72	24.44
Iowa	18.45	22.14	24.91
Kansas	17.32	20.78	23.38
Kentucky	17.30	20.76	23.36
Louisiana	14.63	17.56	19.75
Maine	20.65	24.78	27.88
Maryland	19.60	23.52	26.46
Massachusetts	21.22	25.46	28.65
Michigan	18.71	22.45	25.26

State	Full-time CNA (\$/hr)	PRN CNA (\$/hr)	Agency CNA (\$/hr)
Minnesota	19.40	23.28	26.19
Mississippi	12.35	14.82	16.67
Missouri	15.07	18.08	20.34
Montana	16.83	20.20	22.72
Nebraska	17.00	20.40	22.95
Nevada	19.89	23.87	26.85
New Hampshire	20.36	24.43	27.49
New Jersey	19.02	22.82	25.68
New Mexico	16.20	19.44	21.87
New York	19.56	23.47	26.41
North Carolina	16.24	19.49	21.92
North Dakota	18.33	22.00	24.75
Ohio	16.05	19.26	21.67
Oklahoma	13.39	16.07	18.08
Oregon	18.67	22.40	25.20
Pennsylvania	17.23	20.68	23.26
Rhode Island	19.39	23.27	26.18
South Carolina	15.47	18.56	20.88
South Dakota	14.90	17.88	20.12
Tennessee	15.56	18.67	21.01
Texas	16.94	20.33	22.87
Utah	16.12	19.34	21.76
Vermont	18.60	22.32	25.11
Virginia	17.10	20.52	23.09
Washington	20.80	24.96	28.08
West Virginia	14.22	17.06	19.20
Wisconsin	17.58	21.10	23.73
Wyoming	17.38	20.86	23.46

Table A13. ICD-10-CM Chapters and corresponding categories.

Chapter	Title	Categories
1	Certain Infectious and Parasitic Diseases	A00-B99
2	Neoplasms	C00-D49
3	Diseases of the Blood and Blood-Forming Organs	D50-D89
4	Endocrine, Nutritional, and Metabolic Diseases	E00-E89
5	Mental, Behavioral, and Neurodevelopmental Disorders	F01-F99
6	Diseases of the Nervous System	G00-G99
7	Diseases of the Eye and Adnexa	H00-H59
8	Diseases of the Ear and Mastoid Process	H60-H95
9	Diseases of the Circulatory System	I00-I99
10	Diseases of the Respiratory System	J00-J99

Chapter	Title	Categories
11	Diseases of the Digestive System	K00-K95
12	Diseases of the Skin and Subcutaneous Tissue	L00-L99
13	Diseases of the Musculoskeletal System and Connective Tissue	M00-M99
14	Diseases of the Genitourinary System	N00-N99
15	Pregnancy, Childbirth, and the Puerperium	O00-O9A
16	Certain Conditions Originating in the Perinatal Period	P00-P96
17	Congenital Malformations, Deformations, and Chromosomal Abnormalities	Q00-Q99
18	Symptoms, Signs, and Abnormal Clinical and Laboratory Findings	R00-R99
19	Injury, Poisoning, and Certain Other Consequences of External Causes	S00-T88
20	External Causes of Morbidity	V00-Y99
21	Factors Influencing Health Status and Contact with Health Services	Z00-Z99
22	Codes for Special Purposes	U00-U85, U89

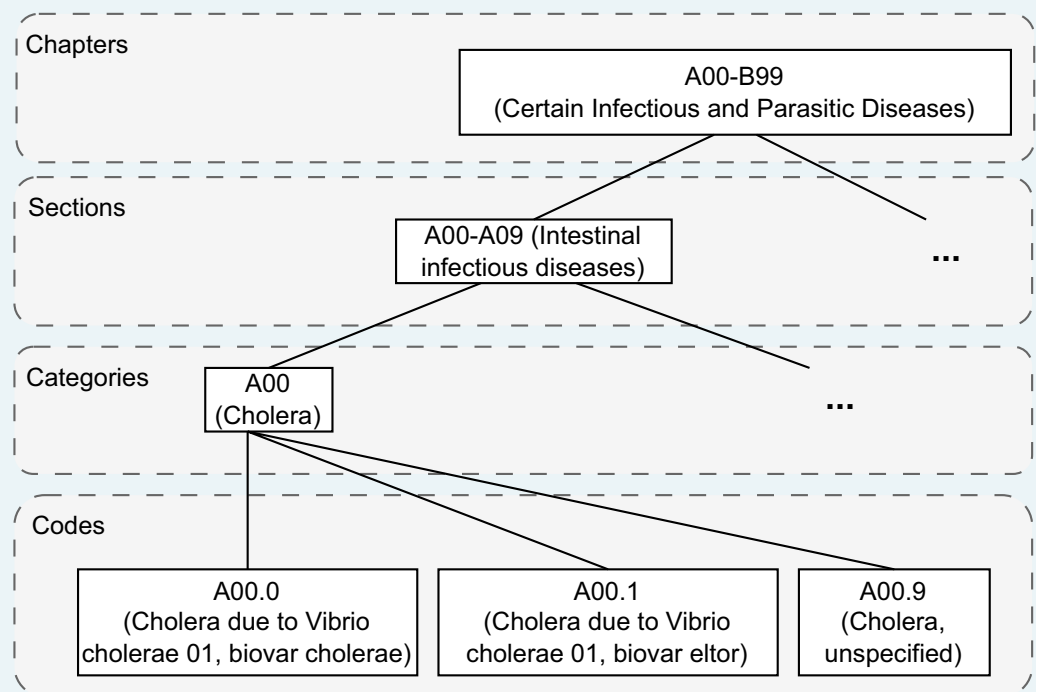


Figure A2. Example of the ICD-10-CM hierarchy for a specific cholera diagnosis.

This hierarchy tree demonstrates that a cholera diagnosis falls under chapter 1 (certain infectious and parasitic diseases), section A00-A09 (intestinal infectious diseases), category A00 (cholera), and could ultimately be one of codes A00.0, A00.1, or A00.9 (biovar cholerae, biovar eltor, or unspecified). Beginning from the top layer and descending, there is an increase in group specificity.

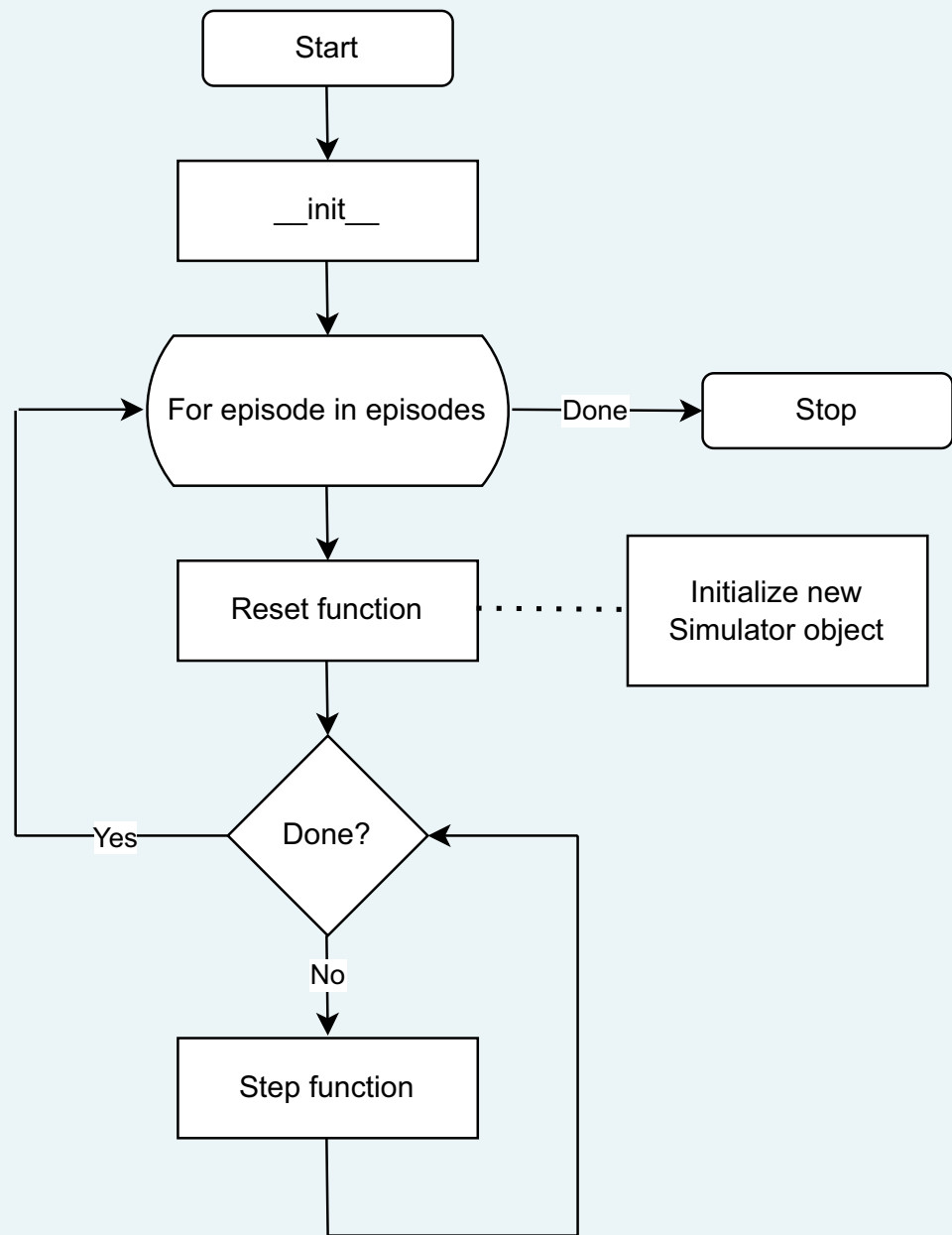


Figure A3. Flow of environment for the snf_v0 Gymnasium environment.

Table A14. PPO Hyperparameters

Hyperparameter	Value
Learning rate (α)	3×10^{-4}
Batch size	64
Steps per update	2048
Epochs per batch	10
Clipping parameter (ϵ)	0.2
Max gradient norm	0.5
Discount factor (γ)	0.99
GAE λ	0.95

Hyperparameter	Value
Entropy coefficient	0.01
Total timesteps	200,000
Policy architecture	MultiInputPolicy