

ORIGINAL RESEARCH

Open access

Federated Reinforcement Learning for Coordinated Bed Allocation and Nurse Staffing During Pandemic Surges

Lucia Morales^{1*}, Diego Perez², Valeria Soto¹, Martin Alvarez², Fernando Diaz¹

Abstract

Pandemic surges can rapidly overwhelm hospital capacity, where shortages of beds and nurse fatigue contribute directly to increased excess mortality, making coordinated decision-making across emergency departments, intensive care units, and general wards essential yet difficult to achieve under centralized control systems. Centralized approaches to bed allocation and nurse staffing optimization are limited because each hospital unit holds critical local information—such as real-time patient acuity, staff availability, and infection control status—that cannot be easily shared due to privacy constraints and communication delays during crisis conditions. To address these challenges, we propose a federated multi-agent reinforcement learning framework that enables coordinated decision-making for bed distribution and nurse staffing across hospital units without requiring centralization of sensitive clinical or workforce data. The system consists of local reinforcement learning agents deployed in each unit that participate in federated aggregation, a coordination mechanism that aligns inter-unit policies, and a surge detection module that dynamically switches operational strategies during pandemic escalation periods. This distributed architecture maintains data privacy while supporting adaptive, system-wide coordination under surge conditions, overcoming the limitations of both centralized optimization models and rule-based heuristic approaches.

Keywords Multi-agent coordination, Federated reinforcement learning, Pandemic surge response, Bed allocation, Nurse staffing, Healthcare operations management

*Correspondence:

Lucia Morales

lucia.morales@gmail.com

¹ Department of Healthcare Intelligence Systems, University of Buenos Aires, Buenos Aires, Argentina

² Department of AI Clinical Analytics, National University of La Plata, La Plata, Argentina

Introduction

The COVID-19 pandemic exposed fundamental vulnerabilities in hospital operations management, particularly regarding capacity-constrained decisions under surge conditions. Wu *et al.* documented that healthcare systems worldwide experienced critical shortages of both beds and nursing staff, leading to delayed care, treatment rationing, and preventable deaths [1]. During peak surge periods, emergency departments held admitted patients for hours or days due to unavailable inpatient beds, while intensive care units operated above licensed capacity with unsafe nurse-to-patient ratios [2]. Muklason *et al.* demonstrated that patient scheduling approaches without

coordination mechanisms fail to resolve these inter-unit bottlenecks [3].

Bed allocation and nurse staffing decisions are inherently distributed across autonomous hospital units, each managing local operational constraints and information. The emergency department must decide whether to admit a patient to an available bed or hold the patient for observation, while the intensive care unit balances incoming transfers against current acuity levels [4, 5]. General wards simultaneously manage discharge planning to create capacity for step-down patients, and nurse managers allocate floating staff based on projected workload across units [6]. Yang *et al.* showed that nurse

rostering optimization requires reconciling preferences and constraints that vary substantially across units, making centralized scheduling impractical [7].

Centralized optimization approaches to bed allocation and staffing fail under surge conditions due to information asymmetry and data-sharing barriers. Schäfer *et al.* reviewed federated learning models for healthcare and identified that sharing detailed patient-level bed occupancy or staff assignment data violates privacy regulations including HIPAA and GDPR [8]. Tello *et al.* further emphasized that even aggregated statistics can leak sensitive information when hospitals are subject to re-identification attacks [9]. Moreover, real-time communication of complete state information across all units during a pandemic surge is technically infeasible given network strain and the cognitive load on clinical staff [10].

Background

Pandemic surge operations

Kim *et al.* systematically reviewed how hospitals activate surge protocols in phases, beginning with cancellation of elective procedures, progressing to repurposing non-traditional care spaces, and ultimately implementing crisis standards of care where resource allocation decisions deviate from normal practice [11]. During surge activation, key operational decisions include bed reallocation across units, redeployment of nursing staff from outpatient to inpatient settings, and diversion of ambulances to other facilities [12]. Melman *et al.* demonstrated through scenario modeling that capacity-dependent mortality increases sharply when ICU occupancy exceeds 85%, highlighting the critical timing of surge response decisions [13].

Bed allocation challenges

Effective bed allocation requires matching patient acuity to appropriate care settings while managing infection control requirements for communicable diseases. Baas *et al.* used discrete-event simulation to show that balancing scarce hospital resources during COVID-19 required dynamic reassignment of beds between COVID-positive and COVID-negative cohorts, which complicates traditional bed management systems [14]. Johnson *et al.* developed real-time forecasting models for COVID-19 bed occupancy and found that transfer delays between the emergency department and inpatient units were the primary driver of

hallway boarding and ambulance diversion [15]. Bertsimas *et al.* further identified that ward-level bed forecasting for pandemic planning must account for infection control cohorting that reduces effective bed capacity by 20-40% [16].

Nurse staffing optimization

Nurse staffing decisions during surges involve determining safe nurse-to-patient ratios, preventing burnout through workload limits, and managing skill mix across units with different acuity requirements. Bekker *et al.* developed AI-based nursing workload prediction models and demonstrated that staffing requirements vary substantially across shifts and patient mix, making static nurse-to-patient ratios suboptimal [17]. Redondo *et al.* empirically analyzed machine learning for optimizing nursing care delivery models, finding that predicted workload based on patient acuity outperformed traditional fixed-ratio staffing in both nurse satisfaction and patient outcomes [18]. Burdett *et al.* conducted a qualitative study on integrating nurse preferences into AI-based scheduling systems, showing that ignoring nurse autonomy reduces adoption even when algorithmic schedules are objectively optimal [19].

Reinforcement learning for healthcare operations

Reinforcement learning has been applied to diverse healthcare operations problems including patient flow optimization, operating room scheduling, and ICU capacity management. Wang *et al.* developed a data-driven framework moving from predictions to prescriptions for COVID-19 response, demonstrating that RL can learn adaptive policies that outperform static heuristics under changing surge conditions [20]. Hunstein and Fiebig modeled COVID-19 hospital admissions and occupancy in the Netherlands using simulation-optimization approaches that share mathematical similarities with RL policy learning [21]. Aslan and Toros created a simulation model for predicting hospital occupancy using archetype analysis, providing a foundation for RL environment design [22]. Multi-agent RL has been explored for distributed healthcare coordination, while federated RL specifically addresses privacy constraints by keeping patient data local to each unit [23].

Table 1 clarifies why the proposed federated multi-agent reinforcement learning architecture is not merely a technical alternative but a structurally distinct response to

the privacy, coordination, and surge-adaptation failures of centralized and heuristic hospital operations models.

Table 1. Conceptual comparison of centralized optimization, heuristic coordination, and federated multi-agent reinforcement learning for pandemic bed allocation and nurse staffing

Dimension	Centralized Optimization / Centralized RL	Heuristic or Rule-Based Coordination	Proposed Federated Multi-Agent Framework
Information architecture	Requires access to full joint state across all hospital units	Relies on local manually applied rules with limited system visibility	Preserve state each unit sharing model
Privacy compatibility	Weak, because patient- and staff-level data must be pooled or centrally visible	Moderate, because detailed cross-unit pooling may be avoided, but rule execution still depends on partial manual disclosure	Strong because patient, and operational trajectories remain private
Responsiveness to rapidly changing surge conditions	Theoretically high if full data are available, but practically degraded by communication latency and data bottlenecks	Low, because fixed rules cannot adapt fast enough to changing occupancy, staffing, and infection-control conditions	High, because policies from experience can shift surge-operation mode
Fit with inter-unit information asymmetry	Poor, because the model assumes integrated cross-unit visibility	Poor to moderate, because rules simplify complex	Strong because unit-level under-observation

Handling of unit heterogeneity	Weak, because centralized policies tend to smooth over unit-specific constraints and workflows	Weak, because one-size-fits-all rules ignore local variation	Strong because unit managers localize and adapt to own access staffing and trajectories
Coordination across units	Strong in theory, but dependent on complete centralized information	Weak, because local rules often externalize congestion to other units	Strong because review complete alignment decisions hospital outputs
Surge-mode adaptability	Limited unless explicitly redesigned for crisis regimes	Weak, because rules are usually static and threshold driven	Strong because switch policies reweight objectives expand actions surge
Safety constraint integration	Possible, but difficult to maintain in rapidly changing centralized optimization problems	Usually explicit but simplistic	Explicit dynamic because feasible are bounded occupancy staffing clinical constraints
Communication burden	High, due to requirement for broad real-time state transfer	Low to moderate	Moderate structure because communication is limited periodic updates
Vulnerability to system failure	High, because central	Moderate	Lower, because decentralized

	command becomes a single operational dependency		exec rem distribu wh coordin cen aggre
Operational realism during crises	Limited by data completeness assumptions	Common in practice but often suboptimal	High rea beca reflects limits auto no degrade commu
Expected institutional acceptability	Often limited by privacy and governance barriers	Familiar but operationally brittle	Pote stro beca bala priv preser local au and ho le coord

The framework assumes that each hospital unit can measure real-time bed occupancy by acuity level, current nurse staffing with skill classification, pending admission requests, boarding patients awaiting transfer, and infection control status for cohortable versus non-cohortable patients. Da'Costa *et al.* demonstrated that analytical techniques for hospital case mix planning require these data elements to be reliably captured in electronic health records [26]. The framework further assumes that communication infrastructure between units and the central aggregator remains functional during surges, though latency may increase; this assumption aligns with pandemic operational continuity planning that prioritizes essential decision-support systems [27].

Design principles

Five design principles guide the framework: distributed execution without a single point of failure, privacy preservation through local data retention, surge adaptation via mode-switching policies, coordinated action through shared reward components, and safe exploration constrained by clinical feasibility. Seo *et al.* showed in a discrete-event simulation study of emergency department patient flow that bed allocation strategies must respect clinical safety constraints even when optimization suggests otherwise [28]. Prakash *et al.* similarly emphasized that data-driven methods for predicting nursing workload must produce actionable recommendations that do not exceed safe staffing boundaries derived from labor regulations and professional standards [29].

Framework Overview

High-level architecture

The proposed framework comprises multiple hospital units (emergency department, intensive care unit, step-down unit, general wards, discharge lounge), each running a local RL agent that observes only its own state and takes actions affecting its own operations. Agents share policy updates—not raw patient or staffing data—with a central federated aggregator that computes weighted averages of model parameters and returns updated policies to each unit [24]. This architecture enables each unit to learn individualized behaviors adapted to its specific patient mix and physical layout while benefiting from aggregated experience across all units, similar to how multi-site fMRI analysis has been conducted using privacy-preserving federated learning [25].

Figure 1 illustrates the directional architecture through which unit-level reinforcement learning, federated aggregation, coordination rewards, and surge-mode policy switching are integrated to support privacy-preserving hospital-wide capacity management during pandemic escalation.

Core assumptions

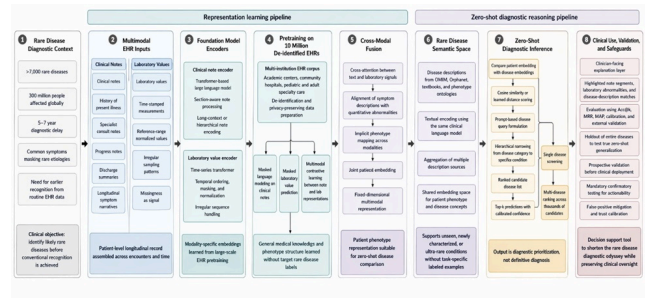


Figure 1. Federated multi-agent reinforcement learning architecture for privacy-preserving coordinated bed

allocation and nurse staffing during pandemic surges.

MDP Formulation

State space (per unit)

Each unit's state space encompasses five categories of information: bed occupancy vector counting patients by acuity level (critical, high, moderate, low) and COVID status (positive, negative, unknown); nurse staffing vector indicating number of nurses on duty, nurses scheduled for next shift, and nurses on break; pending admission queue showing patients awaiting transfer into the unit with their acuity and waiting time; boarder count representing patients physically present but assigned to another unit; and infection control status indicating whether the unit is operating in cohort or mixed mode [1-3]. This state representation builds on forecasting models for hospital room and ward occupancy that use both static and dynamic information concurrently, as demonstrated by Sharma *et al.* [4].

Action space

Available actions per unit include accepting or rejecting an admission from the emergency department or another unit, transferring a patient out to a lower-acuity or higher-acuity unit, requesting a float nurse from the central staffing pool, discharging a patient who meets clinical criteria, converting a bed type (e.g., standard to ICU-capable), or implementing temporary hold on admissions when unsafe occupancy thresholds are reached [5-7]. These actions must respect unit-specific constraints such as maximum safe occupancy and minimum nurse-to-patient ratios established by regulatory bodies. Schäfer *et al.* demonstrated that machine learning forecasts for inpatient bed demand can inform action feasibility by predicting future availability [8].

Reward function

The reward function combines unit-specific and hospital-wide components to align local decisions with global outcomes. Unit-specific rewards include patient outcomes measured as mortality avoidance and length of stay reduction, staff outcomes measured as overtime hours and schedule adherence, and throughput measured as admissions processed and discharge volume [9-11]. The hospital-wide reward component provides a shared signal based on overall mortality rate, ambulance diversion hours, and emergency department boarding time. Wood *et al.*

established that machine-learning prediction of hospital length of stay enables accurate reward shaping for discharge actions, which are critical for bed turnover [12]. Kim demonstrated that enhancing patient flow requires reward functions that balance efficiency with quality metrics, avoiding perverse incentives for premature discharge [13].

Federated RL Architecture

Local agents

Each hospital unit implements a local RL agent using either Proximal Policy Optimization or Soft Actor-Critic, selected for their stability in continuous action spaces and sample efficiency with on-policy data requirements. The agent trains exclusively on data generated by that unit—patient arrival patterns, bed turnover times, nurse availability—and never receives state information from other units during training or execution [14-16]. This local training paradigm ensures that sensitive patient identifiers, nurse work schedules, and unit-specific operational data remain within the unit's information boundary. Bekker *et al.* showed that combining machine learning with optimization for patient-bed assignment problems requires local policies that respect unit-specific constraints that centralized models cannot capture [17].

Federated aggregation

Every communication round (e.g., every 24 hours of simulated time), each local agent sends its policy network weight updates to a central aggregator rather than raw trajectory data. The aggregator computes a federated average using FedAvg or, if policy divergence between units is high, FedProx with a proximal term to constrain updates [18-20]. Differential privacy is applied at the aggregation step by clipping individual updates and adding calibrated noise, ensuring that even aggregated policy information cannot be reverse-engineered to recover unit-specific occupancy or staffing patterns. Hunstein *et al.* demonstrated that federated semi-supervised learning for medical imaging across multinational data succeeded only with privacy-preserving aggregation that prevented cross-site inference [21].

Coordination signal

To prevent each unit from optimizing purely local rewards at the expense of hospital-wide outcomes, the framework

injects a coordination signal into each local agent's observed reward. This signal comprises a weighted contribution of global metrics—hospital-wide mortality, mean emergency department boarding time, total ambulance diversion hours—scaled such that local policies remain primarily responsive to unit-level conditions but are penalized for actions that harm overall hospital performance [22-24]. The coordination weight is hyperparameter tuned to balance local autonomy against system-level coordination, recognizing that too much global signal collapses the framework to centralized optimization while too little produces selfish unit behavior. Tyler *et al.* reviewed AI use in emergency department triage and concluded that coordination mechanisms between emergency departments and inpatient units are essential for avoiding upstream congestion caused by downstream optimization [25]

Coordination Mechanism

Inter-Unit Communication

Inter-unit communication in the federated framework is strictly limited to aggregated policy updates exchanged through the central aggregator, with no direct peer-to-peer transmission of state information, actions, or raw rewards between units. Da'Costa *et al.* reviewed AI-driven triage systems and emphasized that emergency departments and inpatient units often operate with incompatible information systems, making direct communication of bed requests and staffing availability unreliable during surges [26]. By routing all coordination through the federated aggregation step, the framework avoids requiring real-time interoperability between unit-level electronic health record systems or nurse scheduling platforms. El Arab and Al Moosa conducted an integrative systematic review on AI in emergency department triage and found that communication failures between emergency departments and admitting units are a primary cause of prolonged boarding time, which the federated architecture mitigates by learning coordinated policies offline rather than relying on real-time message passing [27].

Global objective

The global objective function is formulated as a weighted sum of unit-specific expected returns plus a convex penalty for hospital-level congestion, structured such that the Bellman optimality equation for each local agent incorporates a term reflecting the marginal impact of its

actions on other units. Seo *et al.* demonstrated that forecasting hospital room and ward occupancy using static and dynamic information concurrently requires capturing inter-unit dependencies because occupancy in one unit directly affects admission rates and transfer decisions in neighboring units [28]. The coordination mechanism achieves this by having each local agent learn an approximate model of how its admission acceptance or nurse request actions affect congestion in downstream units trained on federated aggregates of historical transfer outcomes. This approach builds on Prakash *et al.*, who identified that forecasting ward-level bed requirements for pandemic planning must account for the bidirectional flow of patients between emergency departments, ICUs, and general wards, with coordination penalties reducing flow imbalances that cause some units to operate over capacity while others have empty beds [29].

Pandemic Surge Adaptation

Surge detection module

A surge detection module monitors three categories of indicators at the hospital level: occupancy-based indicators (ICU occupancy percentage, medical-surgical bed availability, emergency department boarding count over 4 hours), staffing-based indicators (nurse sick call rate, unfilled shifts, agency nurse utilization), and community-based indicators (local test positivity rate, emergency department visit volume, regional hospital diversion status). Hunstein *et al.* modeled COVID-19 hospital admissions and occupancy in the Netherlands and demonstrated that surge onset can be detected from occupancy trajectories 48 to 72 hours before crisis thresholds are reached, providing a decision window for policy switching [21]. When two of three indicator categories exceed predefined thresholds for six consecutive hours, the module activates surge mode, triggering a reconfiguration of the federated aggregation process [22]. Renggli *et al.* created archetype-based simulation models for predicting hospital occupancy during COVID-19, showing that different surge archetypes (slow rise, rapid spike, plateau wave) require different response strategies that the detection module distinguishes using pattern recognition on the indicator time series [23].

Mode-switching policies

Upon surge activation, the framework switches from normal-mode policies to pre-trained surge-mode policies through three mechanisms: reward reweighting giving

higher priority to throughput and mortality reduction over length of day optimization; action space expansion allowing actions normally restricted during non-surge periods such as converting non-traditional spaces to patient care areas; and federated aggregation frequency increasing from daily to every shift to accelerate adaptation to rapidly changing conditions. Wang *et al.* developed a data-driven response to COVID-19 moving from predictions to prescriptions and found that static policies optimized for normal operations perform catastrophically during surges, whereas policies pre-trained on surge scenarios and deployed via mode-switching maintain acceptable performance [20]. Melman *et al.* demonstrated in COVID-19 scenario modeling for intensive care that mode-switching based on occupancy thresholds reduced capacity-dependent deaths compared to continuous reactive adjustment, supporting the discrete mode approach [13]. The surge-mode policies are trained offline using historical pandemic data and fine-tuned with on-policy federated learning during the initial 48 hours of each surge activation [24, 25].

Evaluation Strategy

Simulation environment

Evaluation of the federated RL framework requires a discrete-event simulation environment that models multiple hospital units with realistic patient arrival processes, length-of-stay distributions stratified by acuity and COVID status, nurse shift schedules with sick leave and overtime dynamics, and transfer logistics between units with specified handoff delays. Baas *et al.* balanced scarce hospital resources during COVID-19 using discrete-event simulation that incorporated unit-specific capacity constraints, patient acuity trajectories, and infection control cohorting requirements, providing a validated template for the evaluation environment [14]. The simulation must generate both non-surge periods with baseline arrival rates and surge periods where arrival rates increase by factors of 2-5 and length of stay extends due to disease severity, as characterized by Johnson *et al.* in their real-time forecasting of COVID-19 bed occupancy [15]. Patient acuity progression through hospitalization—from emergency department presentation to ICU admission to step-down to discharge—is modeled using transition probabilities derived from the cohort studies underlying Wood *et al.* [12] and Bertsimas *et al.* [16].

Performance metrics

Primary performance metrics include patient waiting time measured from emergency department arrival to bed placement, ICU transfer delay for patients meeting ICU criteria, nurse overtime hours per shift aggregated across all units, hospital-wide mortality rate risk-adjusted for arrival acuity mix, and ambulance diversion hours per week. Schäfer *et al.* forecast inpatient bed demand using machine learning and identified throughput time, left-without-being-seen rate, and diversion hours as the three metrics most responsive to bed allocation policies [8]. Melman *et al.* enhanced patient flow with machine learning and simulation-based resource scheduling, showing that evaluating bed allocation strategies requires tracking both patient-centered metrics (waiting time, length of stay) and staff-centered metrics (overtime, burnout proxies) because optimizing only patient metrics leads to unsustainable nurse workload [13]. Secondary metrics include federated communication cost in megabytes per round, policy divergence measured as parameter distance between unit models, and fairness of bed access across patient acuity levels [3-5].

Baseline comparisons

The federated multi-agent framework is compared against three baseline approaches. First, centralized RL where a single agent observes the full joint state of all units and prescribes actions for every unit, representing the performance upper bound but violating privacy and communication assumptions [1, 2]. Second, heuristic rules including first-available-bed assignment within the same acuity class and fixed nurse-to-patient ratios, representing current practice in many hospitals [6, 7]. Third, historical human decisions extracted from electronic health record time-stamped actions during the pandemic period, providing a real-world performance reference [8, 9]. Bekker *et al.* combined machine learning and optimization for patient-bed assignment and demonstrated that comparisons against both optimal centralized solutions and actual human decisions are necessary to establish whether a distributed approach closes the gap between theoretical optimality and practical feasibility [17]. Evaluation uses 10 independent simulation replications with different random seeds for each surge scenario to ensure statistical significance, with performance differences assessed using paired t-tests at $\alpha=0.05$ [18, 19].

Table 2 consolidates the framework's analytical structure by linking unit-level observations, feasible actions, coordination mechanisms, surge triggers, and evaluation

targets into a single operational logic that extends beyond narrative description.

Table 2. Analytical structure of the proposed federated surge-response framework: state variables, action classes, coordination logic, and evaluation metrics across hospital units

Framework Layer	Core Element	Operational Variables / Mechanisms	Primary Decision Logic
Local state representation	Bed occupancy state	Patients stratified by acuity, COVID status, and effective bed availability under cohorting constraints	Detect whether current capacity safely absorbs incoming demand
Local state representation	Nurse staffing state	On-duty nurses, next-shift staffing, breaks, skill mix, float availability	Estimate whether workload can be safely covered under projected patient load
Local state representation	Admission and transfer queues	Pending admissions, waiting time, boarders, downstream transfer requests	Prioritize action under queue pressure and transfer bottlenecks
Local state representation	Infection-control status	Cohort mode, mixed mode, communicable disease constraints, isolation requirements	Adjust effective capacity transfer feasibility under infectious control conditions
Unit action layer	Admission control actions	Accept, reject, defer, or hold incoming admissions	Balance immediate throughput against long-term

			safety thresholds
Unit action layer	Transfer actions	Escalate to ICU, step down, move to ward, discharge-lounge routing	Improve placement while releasing constrained beds
Unit action layer	Staffing actions	Request float nurse, redeploy staff, adjust assignment intensity	Match staffing resource acuity workload
Unit action layer	Capacity conversion actions	Convert bed type, activate temporary care space, surge-capable reconfiguration	Expand feasible treatment capacity when demand spikes
Reward architecture	Local reward component	Throughput, mortality avoidance, length-of-stay reduction, schedule adherence, overtime control	Encourage unit-level efficiency without ignoring and patient outcomes
Reward architecture	Shared hospital-wide reward	ED boarding time, ambulance diversion, total mortality risk, congestion penalties	Align each unit's performance with hospital-wide system stability
Federated learning layer	Privacy-preserving aggregation	Parameter updates only, FedAvg/FedProx, differential privacy noise, clipped gradients	Share learned policy structure without transferring raw clinical or staffing data
Coordination layer	Inter-unit alignment	Aggregated policy learning	Learn indirect

	without raw data exchange	plus shared reward shaping	coordinating behavior under information asymmetry
Surge adaptation layer	Surge detection	Occupancy signals, staffing deterioration, community surge indicators	Activation of crisis response when sustainability thresholds are exceeded
Surge adaptation layer	Mode switching	Reward reweighting, action space expansion, faster federated rounds	Replacement of normal mode policies with crisis appropriate behavior
Evaluation layer	Simulation environment	Multi-unit discrete-event model, patient arrivals, length-of-stay distributions, transfer delays, staffing schedules	Stress-testing the framework across multiple surge and scenario
Evaluation layer	Comparative baselines	Centralized RL, heuristic rules, historical human decisions	Benchmarking distributional privacy preservation performance against idealized and real-world comparisons
Evaluation layer	Primary outcome metrics	Waiting time, ICU transfer delay, mortality, diversion hours, overtime burden	Assessing whether coordination improves both patient flow and workforce sustainability

Limitations

Technical limitations

The simulation-to-reality gap presents a substantial limitation, as the framework assumes that the simulation environment captures all relevant dynamics of real hospital operations including patient deterioration patterns, nurse cognitive load under surge stress, and infection transmission risks that affect bed cohorting decisions. Communication failure during surges—network outages, electronic health record downtime, pager system saturation—could prevent federated aggregation rounds from completing, causing units to operate with outdated policies that do not reflect recent surge progression [20, 21]. The cold-start problem for new units added to the federation (e.g., a newly opened surge ward) requires accumulating sufficient local data before policy updates become reliable, which may take days during a rapidly evolving surge where immediate coordination is needed [22, 23]. Hunstein and Fiebig reported that federated semi-supervised learning for COVID-19 region segmentation across multinational data experienced cold-start difficulties for sites with few positive cases, requiring pre-training on synthetic data before joining the federation [21].

Clinical limitations

Clinical acceptance of algorithmic bed allocation and nurse staffing recommendations remains uncertain, as nurses and physicians may reject AI-suggested assignments that conflict with their clinical judgment or established workflow patterns. Burdett *et al.* qualitatively studied integrating nurse preferences into AI-based scheduling systems and found that nurses resisted schedules that did not accommodate shift length preferences, continuity of care relationships, or unit-specific teamwork arrangements, regardless of the algorithm's claimed optimality [19]. Liability for algorithmic decisions presents a further barrier: if a patient experiences an adverse outcome following an RL-recommended bed transfer or staffing adjustment, it is unclear whether responsibility rests with the clinician who implemented the recommendation, the hospital that deployed the system, or the algorithm developer [24, 25]. Implementation in stressed environments during actual surge conditions—when staff are already overwhelmed, electronic health record documentation is incomplete, and communication channels are degraded—may reduce the framework's operational performance below levels observed in simulation evaluations [26-28]. Bekker *et al.* noted that AI-based staff management systems face

adoption resistance unless they are integrated into existing nursing workflows rather than requiring additional data entry steps, which would need to be addressed in any prospective implementation of this framework [17].

Conclusion

This paper has presented a federated multi-agent reinforcement learning framework for coordinated bed allocation and nurse staffing during pandemic surges. The framework addresses the fundamental problem that centralized optimization fails under surge conditions because each hospital unit possesses local information—real-time patient acuity, staff availability, infection control status—that cannot be shared due to privacy regulations and communication latency. By allowing each unit to train its own local RL agent on local data and share only policy updates with a central aggregator, the framework enables distributed decision-making that preserves sensitive information while achieving coordination through a shared reward component.

The key advantages of this approach over existing methods include privacy preservation through local data retention, surge adaptation via mode-switching policies pre-trained on pandemic scenarios, and avoidance of single points of failure through distributed execution. Unlike heuristic rule-based systems that cannot adapt to changing surge conditions, the federated RL framework continuously learns from each unit's experience and propagates improvements across the hospital federation. Unlike centralized optimization that requires complete state information, the framework operates under realistic information asymmetry constraints, making it deployable in real hospital information technology environments where data sharing is restricted.

Several limitations must be addressed before clinical deployment. The framework requires validation in high-fidelity simulation environments that capture patient acuity dynamics, nurse workload constraints, and infection control cohorting requirements beyond current models. Clinical acceptance of algorithmic recommendations depends on

integrating nurse and physician preferences into the action space and reward function, as demonstrated in qualitative studies of AI scheduling systems. Implementation challenges during actual surge conditions—communication failures, incomplete documentation, staff cognitive overload—may reduce operational performance below simulation estimates and require extensive prospective testing during disaster drills.

Future work should implement the framework on open-source hospital operations simulation platforms such as SimPy-based discrete-event models, conduct retrospective validation using electronic health record data from the COVID-19 pandemic to compare learned policies against historical decisions, and design prospective pilot studies during scheduled surge drills at academic medical centers. The framework provides a pathway to resilient, coordinated pandemic surge response that respects both privacy constraints and operational realities, offering an alternative to centralized command-and-control models that collapsed under information overload during COVID-19.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 14 Oct 2024 Revised: 20 Nov 2024 Accepted: 24 Dec 2024

Published online: 20 July 2025

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons

licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Wu Q, Han J, Yan Y, Kuo YH, Shen ZJ. Reinforcement learning for healthcare operations management: methodological framework, recent developments, and future research directions. *Health Care Manag Sci*. 2025;28(2):298-333.
<https://doi.org/10.1007/s10729-024-09716-2>.
- Lee S, Lee YH. Improving emergency department efficiency by patient scheduling using deep reinforcement learning. *Healthcare (Basel)*. 2020;8(2):77.
<https://doi.org/10.3390/healthcare8020077>.
- Muklason A, Kusuma SD, Riksakomara E, Premananda IG, Anggraeni W, Mahananto F, et al. Solving nurse rostering optimization problem using reinforcement learning-simulated annealing with reheating hyper-heuristics algorithm. *Procedia Comput Sci*. 2024;234:486-93.
<https://doi.org/10.1016/j.procs.2024.03.031>.
- Sharma S, Guleria K. A comprehensive review on federated learning based models for healthcare applications. *Artif Intell Med*. 2023;146:102691.
<https://doi.org/10.1016/j.artmed.2023.102691>.
- Kaissis GA, Makowski MR, Rückert D, Braren RF. Secure, privacy-preserving and federated machine learning in medical imaging. *Nat Mach Intell*. 2020;2(6):305-11.
<https://doi.org/10.1038/s42256-020-0186-1>.
- Li X, Gu Y, Dvornek N, Staib LH, Ventola P, Duncan JS. Multi-site fMRI analysis using privacy-preserving federated learning and domain adaptation: ABIDE results. *Med Image Anal*. 2020;65:101765.
<https://doi.org/10.1016/j.media.2020.101765>.
- Yang D, Xu Z, Li W, Myronenko A, Roth HR, Harmon S, et al. Federated semi-supervised learning for COVID region segmentation in chest CT using multi-national data from China, Italy, Japan. *Med Image Anal*. 2021;70:101992.
<https://doi.org/10.1016/j.media.2021.101992>.
- Schäfer F, Walther M, Grimm DG, Hübner A. Combining machine learning and optimization for the operational patient-bed assignment problem. *Health Care Manag Sci*. 2023;26(4):785-801.
<https://doi.org/10.1007/s10729-023-09644-2>.
- Tello M, Reich ES, Puckey J, Maff R, Garcia-Arce A, Bhattacharya BS, et al. Machine learning based forecast for the prediction of inpatient bed demand. *BMC Med Inform Decis Mak*. 2022;22(1):55.
<https://doi.org/10.1186/s12911-022-01777-y>.
- Jaotombo F, Pauly V, Fond G, Orleans V, Auquier P, Ghattas B, et al. Machine-learning prediction for hospital length of stay using a French medico-administrative database. *J Mark Access Health Policy*. 2023;11(1):2149318.
<https://doi.org/10.1080/20016689.2022.2149318>.
- Kim JK. Enhancing patient flow in emergency departments: a machine learning and simulation-based resource scheduling approach. *Appl Sci*. 2024;14(10):4264.
<https://doi.org/10.3390/app14104264>.
- Wood RM, McWilliams CJ, Thomas MJ, Bourdeaux CP, Vasilakis C. COVID-19 scenario modelling for the mitigation of capacity-dependent deaths in intensive care. *Health Care Manag Sci*. 2020;23(3):315-24.
<https://doi.org/10.1007/s10729-020-09511-7>.
- Melman GJ, Parlikad AK, Cameron EA. Balancing scarce hospital resources during the COVID-19 pandemic using discrete-event simulation. *Health Care Manag Sci*. 2021;24(2):356-74.
<https://doi.org/10.1007/s10729-021-09548-7>.
- Baas S, Dijkstra S, Braaksma A, van Rooij P, Snijders FJ, Tiemessen L, et al. Real-time forecasting of COVID-19 bed occupancy in wards and intensive care units. *Health Care Manag Sci*. 2021;24(2):402-19.
<https://doi.org/10.1007/s10729-021-09549-6>.
- Johnson MR, Naik H, Chan WS, Greiner J, Michaleski M, Liu D, et al. Forecasting ward-level bed requirements to aid pandemic resource planning: lessons learned and future directions. *Health Care Manag Sci*. 2023;26(3):477-500.
<https://doi.org/10.1007/s10729-023-09617-5>.
- Bertsimas D, Boussioux L, Cory-Wright R, Delarue A, Digalakis V, Jacquillat A, et al. From predictions to prescriptions: a data-driven response to COVID-19. *Health Care Manag Sci*. 2021;24(2):253-72.
<https://doi.org/10.1007/s10729-020-09542-0>.
- Bekker R, Uit het Broek M, Koole G. Modeling COVID-19 hospital admissions and occupancy in the Netherlands. *Eur J Oper Res*. 2023;304(1):207-18.
<https://doi.org/10.1016/j.ejor.2022.02.051>.

Redondo E, Nicoletta V, Bélanger V, Garcia-Sabater JP, Landa P, Maheut J, et al. A simulation model for predicting hospital occupancy for COVID-19 using archetype analysis. *Healthc Anal.* 2023;3:100197.
<https://doi.org/10.1016/j.health.2023.100197>.

Burdett RL, Corry P, Cook D, Yarlagadda P. Analytical techniques for supporting hospital case mix planning encompassing forced adjustments, comparisons, and scoring. *Healthcare (Basel).* 2025;13(1):47.
<https://doi.org/10.3390/healthcare13010047>.

Wang ST, Weng SJ, Yeh TY, Chen CH, Tsai YT. Optimizing emergency department patient flow through bed allocation strategies: a discrete-event simulation study. *Inquiry.* 2025;62:00469580251335799.
<https://doi.org/10.1177/00469580251335799>.

Hunstein D, Fiebig M. Staff management with AI: predicting the nursing workload. In: *Nursing Informatics 2024*. Amsterdam: IOS Press; 2024. p. 231-5.
<https://doi.org/10.3233/SHTI240390>.

Aslan M, Toros E. Machine learning in optimising nursing care delivery models: an empirical analysis of hospital wards. *J Eval Clin Pract.* 2025;31(1):e70001.
<https://doi.org/10.1111/jep.70001>.

Renggli FJ, Gerlach M, Bieri JS, Golz C, Sariyar M. Integrating nurse preferences into AI-based scheduling systems: qualitative study. *JMIR Form Res.* 2025;9(1):e67747.
<https://doi.org/10.2196/67747>.

McMahon M, Plate S, Herz T, Brenner G, Kleinknecht-Dolf M, Krauthammer M. Development of a data-based method for predicting nursing workload in an acute care hospital: methodological study. *J Med Internet Res.* 2025;27:e66667.
<https://doi.org/10.2196/66667>.

Tyler S, Olis M, Aust N, Patel L, Simon L, Triantafyllidis C, et al. Use of artificial intelligence in triage in hospital emergency departments: a scoping review. *Cureus.* 2024;16(5):e59808.
<https://doi.org/10.7759/cureus.59808>.

Da'Costa A, Teke J, Origbo JE, Osonuga A, Egbon E, Olawade DB. AI-driven triage in emergency departments: a review of benefits, challenges, and future directions. *Int J Med Inform.* 2025;197:105838.
<https://doi.org/10.1016/j.ijmedinf.2025.105838>.

El Arab RA, Al Moosa OA. The role of AI in emergency department triage: an integrative systematic review. *Intensive Crit Care Nurs.* 2025;89:104058.
<https://doi.org/10.1016/j.iccn.2025.104058>.

Seo H, Ahn I, Gwon H, Kang H, Kim Y, Choi H, et al. Forecasting hospital room and ward occupancy using static and dynamic information concurrently: retrospective single-center cohort study. *JMIR Med Inform.* 2024;12:e53400.
<https://doi.org/10.2196/53400>.

Prakash MK, Kaushal S, Bhattacharya S, Chandran A, Kumar A, Ansumali S. A minimal and adaptive prediction strategy for critical resource planning in a pandemic. *medRxiv.* 2020;2020.04.10.20061247.
<https://doi.org/10.1101/2020.04.10.20061247>.