Journal of Health Informatics

J Health Inform. 2026;18:e1684
DOI: https://doi.org/10.59681/2175-4411.v18.2026.1684

ORIGINAL ARTICLE

# Integrating machine learning demand forecasting and adaptive operational control in inpatient stroke care: an analysis with real hospital data

## Integração de previsão de demanda por aprendizado de máquina e controle operacional adaptativo no cuidado hospitalar ao AVC: uma análise com dados hospitalares reais

## Integración de la predicción de demanda mediante aprendizaje automático y el control operativo adaptativo en la atención hospitalaria al ACV: un análisis con datos hospitalarios reales

Guilherme Coelho[1], Daniel Correa da Silva[2], Maurício Wesley Perroud Júnior[1], Clarimar José Coelho[2] ✉

ORCID: 0000-0003-3045-4057 · 0000-0003-2418-1961 · 0000-0001-5984-9947 · 0000-0002-5163-2986

[1] Faculdade de Ciências Médicas, Universidade Estadual de Campinas (UNICAMP), Campinas, SP, Brasil.

[2] Escola Politécnica e de Artes, Pontifícia Universidade Católica de Goiás (PUC Goiás), Goiânia, GO, Brasil.

ARTICLE INFORMATION

TYPE
Original Article

IDENTIFIER
e1684

DOI
10.59681/2175-4411.v18.2026.1684

RECEIVED
13 June 2026

ACCEPTED
5 November 2026

REVIEW PROCESS
Double Blind Review

HANDLING EDITOR
Juliano de Souza Gaspar

ACCESS
Open, CC BY 4.0

CORRESPONDING AUTHOR
Clarimar José Coelho
clarimarc@gmail.com

LICENSE

This is an open access article published under the Creative Commons CC BY 4.0 license, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited.

CC BY 4.0

## ABSTRACT

**Objective:** To compare capacity allocation policies for inpatient stroke care integrating Random Forest demand forecasting with machine-learning model predictive control (ML-MPC). **Methods:** Retrospective simulation study using 6,224 stroke admissions at a public neurological reference hospital in Southeastern Brazil (2006–2022). A daily series (2015–2022) was split into training (2015–2019), validation (2020) and test (2021–2022). Base capacity was derived from the 90th percentile of the observed daily census. Six policies were simulated over 727 days, with multi-step forecasts generated recursively from information available at each decision instant; uncertainty was assessed by moving-block bootstrap, re-executing all policies on resampled trajectories. **Results:** ML-MPC attained a mean simulated queue of 0.55 patients and the lowest objective-function value, 58% below the robust policy using 71% less extraordinary capacity (708 fewer units), and 46% below a naive fixed-capacity policy; 95% intervals excluded zero for blocks of 14, 30 and 60 days. Random Forest accuracy exceeded a historical-mean benchmark only marginally (MAE 0.994 vs 0.999). **Conclusion:** Within model conditions, how forecasts, uncertainty and capacity constraints entered sequential decisions mattered more than forecast accuracy. Outcomes are simulation estimates; no patient was exposed, the risk term is a non-validated relative index, and objective weights are judgment-based, not monetary.

**Keywords:** Clinical Decision Support Systems; Machine Learning; Hospital Bed Capacity.

## RESUMO

**Objetivo:** Comparar políticas de alocação de capacidade no cuidado hospitalar ao AVC integrando Random Forest e controle preditivo por aprendizado de máquina (ML-MPC). **Métodos:** Simulação retrospectiva com 6.224 internações por AVC em hospital público de referência do Sudeste do Brasil (2006–2022). A série diária (2015–2022) foi particionada em treino (2015–2019), validação (2020) e teste (2021–2022). A capacidade-base derivou do percentil 90 do censo diário observado. Seis políticas foram simuladas em 727 dias, com previsões multi-passo recursivas usando apenas informação disponível ao decidir; a incerteza foi avaliada por bootstrap em blocos móveis. **Resultados:** A ML-MPC alcançou fila média simulada de 0,55 paciente e o menor valor da função objetivo, 58% abaixo da política robusta usando 71% menos capacidade extraordinária (708 unidades a menos), e 46% abaixo de uma política ingênua de capacidade fixa; os intervalos de 95% excluíram zero. A acurácia do Random Forest superou apenas marginalmente a média histórica (MAE 0,994 contra 0,999). **Conclusão:** Nas condições do modelo, o modo como previsão, incerteza e restrições entraram nas decisões sequenciais importou mais que a acurácia preditiva. Os desfechos são estimativas de simulação; nenhum paciente foi exposto, o risco é índice relativo não validado e os pesos são de julgamento, não monetários.

**Descritores:** Sistemas de Apoio a Decisões Clínicas; Aprendizado de Máquina; Número de Leitos em Hospital.

## RESUMEN

**Objetivo:** Comparar políticas de asignación de capacidad en la atención hospitalaria al ACV integrando Random Forest y control predictivo por aprendizaje automático (ML-MPC). **Métodos:** Simulación retrospectiva de 6.224 internaciones por ACV en hospital público de referencia del Sudeste de Brasil (2006–2022). La serie diaria (2015–2022) se dividió en entrenamiento (2015–2019), validación (2020) y prueba (2021–2022). La capacidad base derivó del percentil 90 del censo diario observado. Se simularon seis políticas en 727 días, con predicciones multipaso recursivas sin información futura; la incertidumbre se evaluó por bootstrap en bloques móviles. **Resultados:** ML-MPC alcanzó cola media simulada de 0,55 paciente y el menor valor de la función objetivo, 58% por debajo de la política robusta usando 71% menos capacidad extraordinaria (708 unidades menos), y 46% por debajo de la política ingenua de capacidad fija; los intervalos del 95% excluyeron cero. El Random Forest superó solo marginalmente a la media histórica (MAE 0,994 frente a 0,999). **Conclusión:** En las condiciones del modelo, cómo entraron predicción, incertidumbre y restricciones en las decisiones secuenciales importó más que la precisión predictiva. Los desenlaces son simulados; ningún paciente fue expuesto, el riesgo es índice relativo no validado y los pesos son de juicio, no monetarios.

**Descriptores:** Sistemas de Apoyo a Decisiones Clínicas; Aprendizaje Automático; Número de Camas en Hospital.

**HOW TO CITE THIS ARTICLE:** Coelho G, Correa da Silva D, Perroud Júnior MW, Coelho CJ. Integrating machine learning demand forecasting and adaptive operational control in inpatient stroke care: an analysis with real hospital data. J Health Inform 2026;18:e1684. doi: https://doi.org/10.59681/2175-4411.v18.2026.1684

## 1 INTRODUCTION

Stroke is the leading cause of death and disability in Brazil, responsible for more than 820,000 hospital admissions between 2018 and 2022 [1], and affects over 12 million people globally each year [2]. The global burden of stroke has been rising steadily, with low- and middle-income countries disproportionately affected [3]. Stroke care comprises two complementary needs. The first is acute intervention: rapid recognition, assessment and reperfusion therapy when indicated. Intravenous thrombolysis is generally indicated for eligible patients within 4.5 hours of symptom onset, although selected patients may be treated later on the basis of imaging criteria, and the benefit of alteplase declines as the interval from onset lengthens [4,5]. Thrombectomy may benefit selected patients up to 24 hours from the time they were last known well, a selection criterion distinct from the neurological surveillance that the first 24 hours of admission require. The second need is organized inpatient care: neurological surveillance, management of complications, secondary prevention, early rehabilitation and multidisciplinary follow-up throughout the admission. This second dimension remains essential whether or not reperfusion is indicated, and organized inpatient stroke care is associated with better outcomes [6]. The model developed here is an operational approximation of inpatient service capacity and therefore addresses this second dimension only: it does not represent the interval between symptom onset and thrombolysis, and the simulated queue must not be read as access to reperfusion therapy.

In Brazil, where 80% of stroke patients depend on the public health system (Sistema Único de Saúde, SUS), the operational sustainability of inpatient stroke care is not merely an administrative concern but a clinical imperative with direct consequences for patient survival and functional recovery [7]. Organized inpatient stroke care is recognized as the most effective model for managing stroke admissions [6], yet its delivery depends on capacity management decisions: how many beds to make available, when to mobilize additional nursing or specialist staff, and how to anticipate peaks in demand.

Interest in quantitative methods for hospital capacity management dates to the foundational work of Erlang in the early 20th century, formalized for healthcare settings through queueing theory [8]. These classical frameworks established that systems with arrival rates close to or exceeding service capacity are inherently unstable, a result with direct implications for stroke care services, where arrivals are stochastic and capacity is constrained by specialist staffing and physical beds. Little's Law, a fundamental result in queueing theory [9], provides the basis for empirically calibrating steady-state capacity requirements from observed length of stay and bed counts, as applied in the present study.

Building on this foundation, the application of machine learning (ML) to hospital demand forecasting gained momentum in the 2010s. Sudarshan *et al.* [10] demonstrated that Random Forest achieves a favorable accuracy–interpretability trade-off in emergency department arrival forecasting across multiple horizons. Porto and Fogliatto [11] extended this to eleven emergency departments across three countries, showing that feature engineering with temporal lags and rolling statistics reduces the mean absolute percentage error (MAPE) to between 3% and 18%. Tello *et al.* [12] applied ML-based forecasting specifically to inpatient bed demand, confirming that data-driven models outperform rolling-average benchmarks in weekly predictions. On the optimization side, robust optimization provides principled guarantees against worst-case demand scenarios at moderate resource cost [13,14], and has been adapted to hospital bed allocation in several studies [15]. Model predictive control (MPC), formalized by Camacho and Bordons [16], integrates short-term forecasts directly into a multi-period cost optimization through a receding-horizon mechanism, and has been applied in biomedical settings, for example in closed-loop insulin delivery [17].

Despite this body of work, a gap persists: forecasting studies rarely translate their outputs into dynamic resource allocation decisions, and optimization studies rarely evaluate on real longitudinal data with empirically calibrated parameters. Their integration in inpatient stroke care remains largely absent from the literature.

This paper addresses that gap by comparing six capacity allocation policies (baseline, naive fixed capacity, historical mean, Random Forest point forecast, robust, and ML-MPC) in a simulation calibrated from 6,224 real stroke admissions at a public neurological reference hospital in Southeastern Brazil, covering 16.7 years (April 2006 to December 2022) [18]. The same admission records were previously analyzed by the authors in an epidemiological study of trends in hospitalization, in-hospital mortality and thrombolysis adoption, which characterized the clinical profile of the cohort. The present study asks a different question. It constructs a daily arrival series that the earlier work did not build, calibrates a discrete-time queueing model of inpatient capacity, and reports operational outcomes (queue length, extraordinary capacity use and an objective-function value), none of which appear in the earlier analysis. There is no overlap in research question, statistical method or reported outcome.

The central hypothesis is that operational performance depends not only on the accuracy of demand forecasting, but on how that forecast, its uncertainty and the capacity constraints are incorporated into

sequential allocation decisions. The contributions are: (i) a comparison of six progressively more elaborate policies on the same real series, including naive comparators that separate the effect of merely having extraordinary capacity available from that of allocating it predictively; (ii) calibration of effective service capacity from the observed daily census, in a service that holds no beds exclusively dedicated to stroke; and (iii) an assessment of how far the resulting ordering of policies depends on the objective-function weights, on the risk parameters, on the start of the modeling window, and on resampling of the demand trajectory.

## 2 METHODS

### 2.1 Dataset

The dataset analyzed in this study has been previously characterized in an earlier publication [18] in an epidemiological analysis of trends in hospitalization, in-hospital mortality, and thrombolysis adoption in the same cohort. The present work uses the same admissions records as input for an operational research framework, with no overlap in research questions, statistical methods, or reported outcomes; patient-level outcomes are used here only as calibration inputs for the cost function and risk dynamics.

Records from 6,224 stroke admissions at a public neurological reference hospital in Southeastern Brazil, covering April 2006 to December 2022 (16.7 years), were used. Each record contains: stroke type, admission unit (ward or intensive care unit [ICU]), admission and discharge dates, length of stay in days, clinical outcome, and therapeutic approach. The distribution by stroke type was: ischemic stroke (IS): 4,581 cases (73.6%); hemorrhagic stroke (HS): 1,187 (19.1%); and subarachnoid hemorrhage (SAH): 456 (7.3%). In-hospital mortality rates were 19.0% (IS), 39.6% (HS), and 53.1% (SAH) (Table 1).

Although admission records are available from April 2006, the modeling window was restricted to 2015–2022. This was a pragmatic choice rather than one forced by a documented change in registration or service organization: the window provides eight complete calendar years while remaining close to the current configuration of the service, whose annual volume grew by roughly 40% between 2006–2014 and 2015–2022. Because the choice is pragmatic, its influence on the results is tested directly in Section 3.5, where the entire protocol is repeated with training beginning in 2010. The daily series runs from 2015-01-01 to 2022-12-28, the date of the last recorded admission, and therefore contains 2,919 days; the final three days of December 2022 carry no admission record and are not represented. The inter-arrival process was

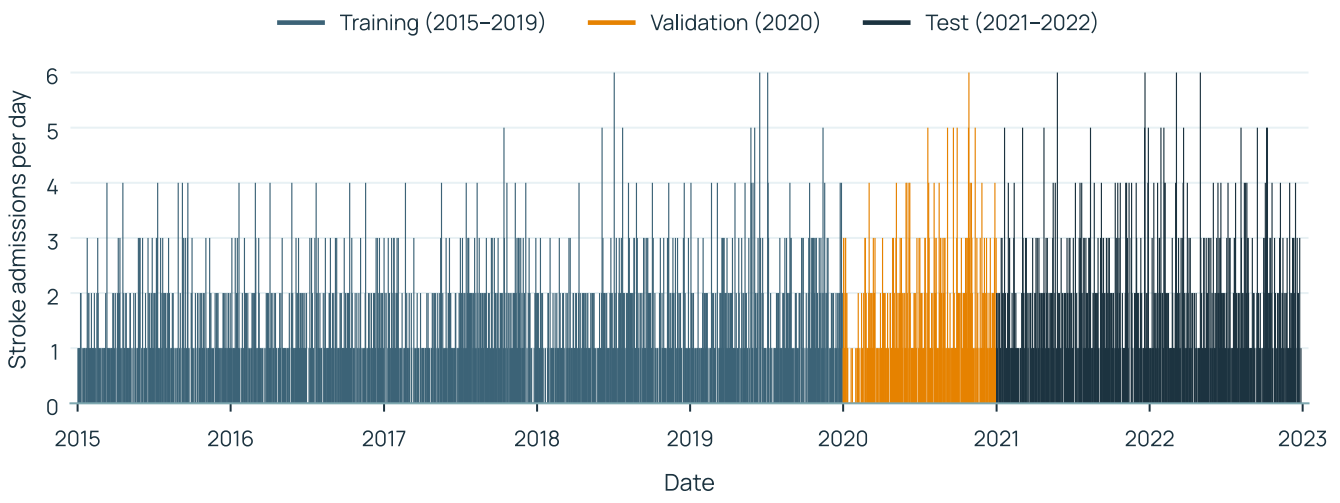
**Table 1.** Clinical profile by stroke type.

| Type | n (%) | Mort. (%) | ICU (%) | LOS (d) |
|---|---|---|---|---|
| IS | 4,581 (73.6) | 19.0 | 4.2 | 9.8 |
| HS | 1,187 (19.1) | 39.6 | 17.1 | 16.2 |
| SAH | 456 (7.3) | 53.1 | 39.3 | 21.6 |
| Total | 6,224 (100) | 25.4 | 9.2 | 11.9 |

IS: ischemic stroke; HS: hemorrhagic stroke; SAH: subarachnoid hemorrhage; Mort.: in-hospital mortality; LOS: mean length of stay.

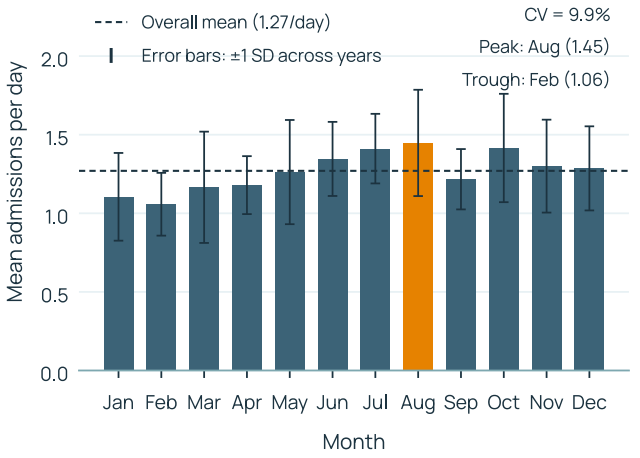
**Source:** Prepared by the authors (2026).

**Figure 1.** Daily time series of stroke admissions (2015–2022), by partition (training, validation and test).




**Source:** Prepared by the authors (2026).

**Figure 2.** Mean daily stroke admissions by calendar month (2015–2022). Monthly coefficient of variation = 9.9%; peak in August, trough in February.




**Source:** Prepared by the authors (2026).

characterized by a variance-to-mean ratio of 1.005, close to the Poisson value of 1.0, and an overall mean of 1.27 admissions per day. The coefficient of variation of the twelve monthly means was 9.9%, indicating moderate seasonality, with a peak in August (1.45 admissions per day) and a trough in February (1.06). Demand grew at an average rate of 5.9% per year. Figure 1 shows the daily series and Figure 2 the monthly pattern.

### 2.2 System Model

The hospital system is modeled as a discrete-time queue with daily granularity. Queue dynamics follow:

$$q_t = \max(0,\ q_{t-1} + a_t - x_t) \tag{1}$$

where qt is the queue at the end of day t, at the observed number of admissions, and xt = b + ut the total service capacity, with base capacity b and extraordinary capacity ut ∈ {0,1,2,3}. Because b is a rate rather than an integer count, the model is a fluid approximation: the number served on day t is st = min(qt−1 + at, xt), a continuous quantity, and qt consequently takes fractional values. These are expected numbers of patients under a deterministic daily balance, not counts of individuals. One unit of ut denotes one additional patient-equivalent of daily throughput (whatever combination of staffing, bed or discharge-pathway measures delivers it) and is not equivalent to one physical bed held open for 24 hours.

Base capacity was calibrated using Little's Law [9] in the form b = L / W, where L is the mean number of patients simultaneously occupying beds (patients), W is the mean length of stay (days), and b is the resulting steady-state discharge rate (patients per day), the rate at which beds become free for new admissions. The hospital does not maintain beds exclusively dedicated to stroke care, so no dedicated bed count was available as an input to the formula. Effective service capacity was therefore estimated independently of the arrival series: admission and discharge dates were used to reconstruct the daily census of stroke inpatients, and the 90th percentile of that census over the modeling window (19 patients) was adopted as the occupancy the service has demonstrably sustained on a recurring basis. With a mean length of stay of 11.9 days, this yields b = 19 patients / 11.9 days ≈ 1.597 discharges per day. Two features of this estimator require explicit statement. First, it is not circular: the mean census (14.6 patients) is the quantity that would reproduce λW and hence return the arrival rate itself, whereas the 90th percentile does not. Second, taking a high percentile of occupancy in place of the mean is an operational assumption about attainable capacity, not a direct application of Little's identity, which relates means. The periods feeding each estimate differ and are stated here in full. The census percentile was computed over 2015–2022 and the mean length of stay over the complete record (2006–2022, Table 1), so both include the 2021–2022 test window. This is a retrospective calibration, admissible in a retrospective simulation, but not equivalent to a prospective deployment with parameters fixed in advance. Three things are kept apart: demand forecasting, which uses no information from at or after the decision instant (Section 2.3); capacity calibration, which is retrospective; and the policy comparison, which is conditional on that calibration. Restricting calibration to the training period gives a census percentile of 20 patients and a mean stay of 11.65 days, hence b = 1.717 per day, a value inside the range swept in Section 3.3, where ML-MPC also attains the lowest objective-function value. Mean demand over the test period is 1.50 admissions per day, giving a utilization of ρ = 0.94 at the adopted calibration; the system is therefore stable, and the baseline policy is a legitimate comparator rather than one that is unstable by construction. Maximum extra capacity $u_{max}$ = 3 corresponds to the difference between the 99th and 75th historical percentiles of daily admissions.

The state variable rt is an operational penalty index that accumulates while patients wait for admission to organized inpatient care and dissipates once they are admitted. It is a construct of the model: it was not validated against any clinical outcome and does not measure neurological deterioration, disability or mortality. Its role is to penalize persistent waiting more heavily than transient waiting when policies are ranked [6]:

$$r_t = \max(0,\ r_{t-1} + \eta\, q_t - \mu\, s_t) \tag{2}$$

where $s_t$ is the number of patients actually served on day t, η governs risk accumulation per patient-day in queue and μ its dissipation per patient served. Because the recursion is positively homogeneous, scaling both η and μ by any factor k scales $r_t$ by exactly k, which is

indistinguishable from scaling the cost weight β by k. The triple (η, μ, β) therefore carries only two degrees of freedom: the ratio η/μ and the product βη. The values η = 0.08 and μ = 0.06 are a normalization within an equivalence class, not two independently estimated quantities; η = 1, μ = 0.75, β = 0.12 yields numerically identical results. The single substantive parameter is the ratio η/μ = 1.33, set so that risk accumulates while patients wait and dissipates once they are admitted, consistent with the established benefit of organized inpatient stroke care. Accordingly, $r_t$ is interpreted only as a relative index for ranking policies, with no absolute clinical meaning, and its sensitivity across the full plausible range of η/μ is reported in Section 3.3. The multi-period cost function integrates queue, risk, and extra capacity:

$$J = \alpha \sum q_t + \beta \sum r_t + \gamma \sum u_t \quad (3)$$

with weights α = 3.0, β = 1.5, and γ = 8.0. The ordering α < γ reflects that extraordinary capacity mobilizes resources from other units, whereas queueing is absorbed internally. Because J is linear in the weights, the region over which each policy dominates admits a closed form, reported in Section 3.3. To avoid ambiguity about which quantities act where: η and μ enter Equation (2) directly and are the only parameters governing the risk recursion. Type-specific severity weights derived from observed in-hospital mortality (IS = 1.00; HS = 2.08; SAH = 2.80) enter neither Equations (1)–(3) nor the forecasting model. They are computed during preprocessing but excluded from the predictor matrix, which contains only the cyclic calendar encodings, the lagged admissions and the rolling means listed in Section 2.3. No predictor therefore depends on the case mix of days not yet observed. The weights are reported here because they document the clinical heterogeneity of the cohort, not because they act anywhere in the model.

### 2.3 Demand Forecasting Model

The series was split into training (2015-01-01 to 2019-12-31; 1,826 days), validation (2020; 366 days) and test (2021-01-01 to 2022-12-28; 727 days), summing to the 2,919 days of the series. The test partition holds 727 rather than 730 days for the same reason given above: the series ends on the date of the last recorded admission. The year 2020 was held out, and the grounds are stated here with their limits. Although the pandemic produced a documented reduction of 22–35% in stroke admissions internationally [19], no such reduction occurred at the study center: 507 admissions were recorded in 2020 against 502 predicted by the 2015–2022 annual trend fitted excluding that year, a deviation of +1.0%. What changed was the observed length of stay, which fell from 10.4 days in 2018–2019 to 8.7 days in 2020 (medians 5.0 and 4.0; Mann–Whitney $p < 0.001$), alongside a non-significant shift toward intensive care (8.8% to 11.6%; $p = 0.10$). Since base capacity enters the model as beds divided by mean length of stay, a shift of this size in the denominator would alter b if the number of beds were held constant; whether effective capacity actually changed in 2020 is not established by these data. Two caveats attach. The trend yielding the expected 507 admissions is fitted over 2015–2022 and so includes the test years, making the reasoning retrospective rather than a criterion applicable before the test period; and a difference in length of stay does not by itself make exclusion statistically necessary. Holding 2020 out is a design decision, motivated by the observed difference in length of stay rather than by a demonstrated change in effective capacity. The Random Forest model [20] used 300 trees and maximum depth 6. Features included: cyclic encoding of day of week and month, lagged admissions at 1, 2, 3, 7, 14, and 30 days, rolling means over 7, 14, and 30 days, and a linear trend. Hyperparameters were fixed a priori; no model selection or hyperparameter search was performed, and 2020 served exclusively as an out-of-sample monitoring set rather than a tuning set. Because the controller optimizes over a four-day horizon, forecasts for days t+1 to t+3 are required at the decision instant t, when the arrivals of those days are by definition unknown. Forecasts are therefore produced recursively: the one-step forecast for day t uses observed arrivals up to t−1, and each subsequent step feeds the model's own prediction back into every temporal feature: the lags at 1, 2 and 3 days and the rolling means over 7, 14 and 30 days alike. No value observed at or after the decision instant enters any forecast. Recursive forecasting degrades accuracy only slightly relative to a one-step forecast computed with observed lags (test mean absolute error [MAE] 0.9941 at horizon 1 against 1.0006 at horizon 4), because the model's most influential features are the 14- and 30-day rolling means and the weekly cyclic encodings rather than the short lags.

### 2.4 Capacity Allocation Policies

Six policies were evaluated over the 727 test days. Baseline: ut = 0 in all periods; no adaptive management. Naive fixed capacity: ut = 1 in all periods, without any forecast, isolating the effect of merely having extra capacity available from that of allocating it predictively. Historical mean: allocates extra capacity proportional to the excess of the historical mean by day of week and month over base capacity, without machine learning. Random Forest point forecast: identical rule applied to the Random Forest point forecast. Robust: replaces the point forecast with a high quantile of the empirical historical distribution by day of week and month, following the general principle of provisioning against adverse realizations rather than against the expected value [14]. The 90th percentile was chosen by the authors as the operating point; that specific choice is not prescribed by the cited work. ML-MPC: applies receding-horizon MPC with H = 4 days [16], a sequential

decision problem under uncertainty in the sense of Powell [21], minimizing the objective function using Random Forest forecasts with a 25% buffer; the subproblem is solved by exhaustive enumeration of the $4^4 = 256$ action sequences, and the full specification of the problem solved at each instant (state projection, treatment of the penalty term, absence of terminal conditions and solution method) is given in the repository cited under Data and Materials Availability.

### 2.5 Uncertainty of the Policy Differences

Results derive from a single 727-day trajectory, so differences between policies carry no sampling error of their own. Uncertainty was therefore assessed by moving-block bootstrap at two levels. The first resamples the paired daily contributions to the objective function and yields intervals conditional on the simulated trajectories. The second resamples blocks of the demand series and re-executes all six policies on each resampled trajectory, so that queues, risk and control decisions are recomputed; this is the stronger of the two and is the one reported. Arrivals and their corresponding forecasts are resampled jointly as pairs, preserving within each block the observed association between demand and forecast error; the forecast attached to a decision therefore derives only from data preceding that decision in the original series, so no future information is introduced. At block junctions the multi-step forecasts no longer match the arrivals that follow. This introduces additional forecast error rather than information, and its effect on any particular comparison is not determined in advance, since it depends on the policy and on the trajectory drawn. Queue and risk are initialized once per replicate and propagated continuously, without reset at junctions. Two thousand replicates were drawn for block lengths of 14, 30 and 60 days, and 95% intervals are reported as percentiles. These intervals describe robustness under resampling of the observed joint distribution of arrivals and forecasts; they do not describe a forecasting system retrained on new temporal trajectories.

### 2.6 Ethics, Reporting, and Data Availability

This study was approved by the Research Ethics Committee of the coordinating institution under CAAE 79472324.7.0000.5404 (opinion no. 6.898.395, 20 June 2024). Informed consent was waived (Resolution 466/2012). The study followed the RECORD guideline [22] and the TRIPOD+AI checklist [23]. Data are deposited in an institutional research data repository with a persistent identifier (https://doi.org/10.25824/REDU/4MZO5C). The analytic pipeline (Python 3.x, seed = 42) is publicly available under an MIT license at https://github.com/clarimar/jhi-1685-reproduction and archived at https://doi.org/10.5281/zenodo.23985198.

## 3 RESULTS

### 3.1 Forecasting Model Performance

The Random Forest model achieved a training MAE of 0.721 (root mean square error [RMSE] 0.927), a validation MAE of 0.947 (RMSE 1.203) and a test MAE of 0.994 (RMSE 1.283). The test error is 1.38 times the training error for both metrics, so the model retains capacity it cannot carry to unseen periods; the strict temporal partition offers no way to tune this away without reintroducing the model selection the protocol excludes. It is not without predictive value: against a naive lag-1 forecast on the same days (MAE 1.337) it is 25.6% more accurate. Against stronger benchmarks the margin narrows sharply. A historical mean by day of week and month attains MAE 0.999 and a 30-day rolling mean 1.039, so on mean absolute error the gain over the historical mean is 0.45%, a difference of 0.005 admissions per day that was not tested for significance and should not be read as established. The improvement is more pronounced on RMSE (1.283 against 1.332), suggesting better performance where larger errors are penalized; characterizing the extremes would require an analysis of the error distribution not undertaken here. A seasonal naive forecast performs worst (MAE 1.378). The 30-day and 14-day rolling means, the cyclic weekly encodings, the 7-day rolling mean and the linear trend lead the mean-decrease-in-impurity ranking (Figure 3), which describes how the fitted model allocates splits and is distributed arbitrarily among correlated predictors; it is not evidence that these quantities drive demand.

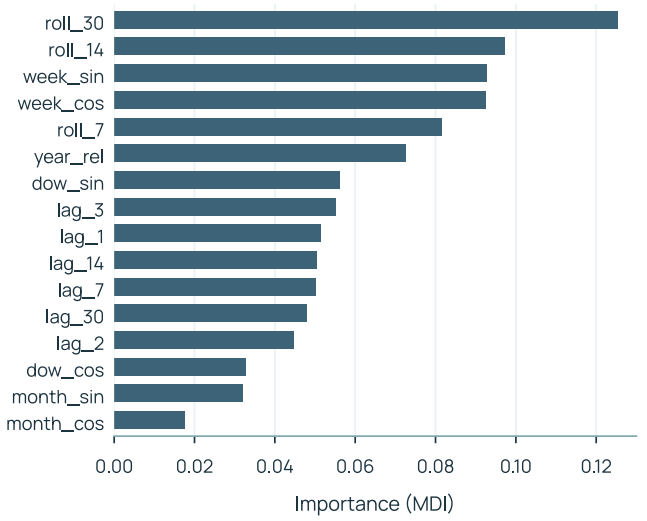
Figure 4 shows the three model-based forecasts against actual daily admissions.

### 3.2 Capacity Policy Comparison

Table 2 presents performance metrics for the six policies over the 727-day test set. The baseline policy absorbs the mean demand of 1.50 admissions per day without divergence, since base capacity is 1.597 discharges per day, but does so at a utilization of 0.94. The simulated queue averages 4.34 patients and peaks at 17.5, with a mean accumulated risk index of 77.4 and an objective-function value of 93,851. Dividing the mean queue by the mean arrival rate gives a wait proxy of 2.90 days. This proxy is a property of the simulated system, not a measurement: no queue for an inpatient bed was observed at the study center, and the figure must not be read as the time patients actually waited. What the result shows is that a service which is formally stable can still accumulate a persistent backlog for organized inpatient care at high utilization.

The naive fixed-capacity policy, which allocates one extra unit every day without any forecast, reduces the mean queue to 0.319 patients and mean risk to 0.016 at a cost of 727 units of extraordinary capacity and

**Figure 3.** Feature importances in the Random Forest model (MDI, mean decrease in impurity).

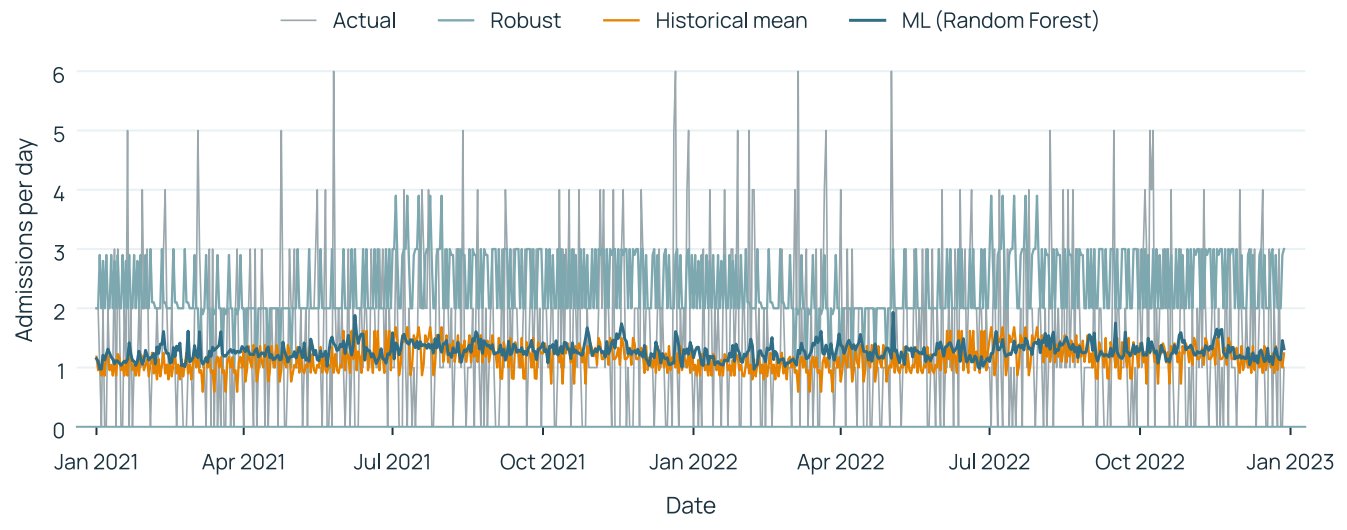



**Source:** Prepared by the authors (2026).

an objective-function value of 6,529 in arbitrary units. This confirms that a substantial share of the achievable gain derives simply from having extraordinary capacity available. The historical-mean and Random Forest point-forecast policies behave close to the baseline (mean queues of 3.71 and 3.81; objective-function values of 72,413 and 77,539), because their point forecasts seldom exceed a base capacity of 1.597 and they therefore allocate extra capacity on very few days (27 and 19 units over 727 days). Accurate point forecasting alone, without explicit treatment of uncertainty, yields little operational benefit.

The robust and ML-MPC policies both hold the simulated system near equilibrium throughout the test window. Their trajectories differ in texture rather than in level: the robust policy allocates extraordinary capacity on most days in small amounts, whereas ML-MPC leaves it at zero for long stretches and raises it in short bursts when the forecast indicates a run of high demand. Figure 5 shows the full trajectories of queue, risk index and extraordinary capacity for all six policies. The robust policy uses 998 units of extraordinary capacity and attains a mean queue of 0.233 patients with an objective-function value of 8,503. ML-MPC uses 290 units (70.9% less, or 708 fewer units of extraordinary capacity over two years) and attains the lowest objective-function value, 3,527, a reduction of 58.5% relative to the robust policy and of 46.0% relative to the naive fixed-capacity policy, whose extraordinary capacity it undercuts by 60.1%. Because the weights of the objective function were set by judgment and do not represent monetary quantities, these reductions are reductions in the value of the objective function and not financial savings. The two policies differ in profile rather than in overall quality. The robust policy keeps the body of the queue distribution lower: its mean queue is 0.233 against 0.546, its 95th percentile 1.40 against 2.40, and it exceeds one patient on 8.7% of days against 21.3%. The robust policy therefore performs better on the frequency and high-percentile metrics of the queue, while ML-MPC records the lower maximum over the simulated trajectory (4.81 against 5.21), and reaches the same order of magnitude of the risk index (0.014 against 0.009) while using less than a third of the extraordinary capacity. No equivalence margin was defined in advance, so these results describe different operating profiles and do not establish equivalence, clinical or statistical, between the two policies. Decomposing the objective function clarifies that the policies are not merely scaled versions of one another. Under the baseline, 89.9% of the value comes from the risk term and 10.1% from the queue term, with no extraordinary capacity used. Under the robust policy

**Figure 4.** Forecasts from the Historical mean, Robust, and ML policies versus actual stroke admissions (2021–2022).




**Source:** Prepared by the authors (2026).

**Table 2.** Policy performance in the test set (2021–2022; 727 days).

| Policy | Qavg | Ravg | Cost | Extra | Wait |
|---|---|---|---|---|---|
| Baseline | 4.342 | 77.378 | 93,851 | 0 | 2.90 |
| Naive (fixed) | 0.319 | 0.016 | 6,529 | 727 | 0.21 |
| Historical mean | 3.713 | 58.779 | 72,413 | 27 | 2.48 |
| RF point forecast | 3.806 | 63.353 | 77,539 | 19 | 2.54 |
| Robust | 0.233 | 0.009 | 8,503 | 998 | 0.16 |
| ML-MPC | 0.546 | 0.014 | 3,527 | 290 | 0.36 |

RF: Random Forest; Qavg: mean simulated queue (patients); Ravg: mean value of the relative risk index, which was not clinically validated; Cost: value of the objective function in abstract units, not currency; Extra: total units of extraordinary capacity; Wait: wait proxy, the mean queue divided by the mean arrival rate (days), a property of the simulation and not an observed waiting time. All statistics refer to the single simulated 727-day trajectory.

**Source:** Prepared by the authors (2026).

the composition inverts: 93.9% comes from extraordinary capacity and 0.1% from risk. Under ML-MPC the value is divided between extraordinary capacity (65.8%) and queue (33.8%), with 0.4% from risk. All policies are evaluated by the same function under the same weights, so the totals remain comparable; what the decomposition shows is that the superiority of ML-MPC in J rests on the structure of the function and on the weights chosen for it, which is precisely the dependence delimited in Section 3.3.

### 3.3 Sensitivity Analyses

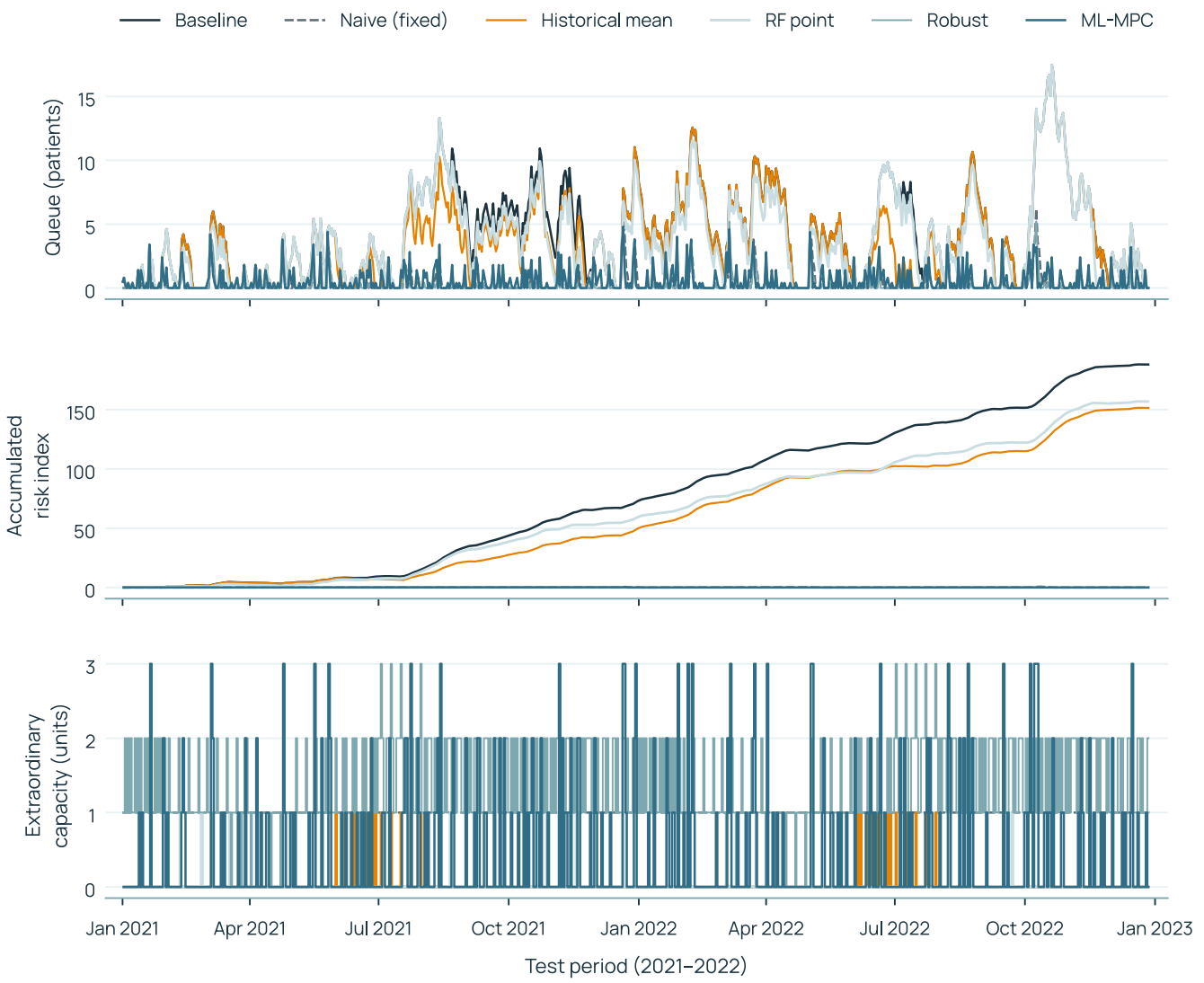
Three analyses delimit the conditions under which the reported ordering holds; the complete grids are available in the repository cited under Data and Materials Availability. Effective capacity was varied from 15 to 23 beds, with and without an occupancy discount of 0.85, spanning utilizations from 0.78 to 1.40 and therefore including unstable configurations: ML-MPC attained the lowest objective-function value in all ten. A 5 × 5 grid over η and μ, spanning ratios from 0.44 to 4.00, placed ML-MPC first in 24 of 25 cells, the exception being the extreme corner η/μ = 4.00; its mean queue varied only between 0.531 and 0.554 across the grid. Prevailing in 24 of 25 cells bounds the sensitivity to η and μ; it does not remove the dependence of the comparison on the form of the objective function itself.

For the cost weights, linearity of J yields a closed-form dominance condition. Writing ΔS for the differences in accumulated sums between ML-MPC and the robust policy (ΔS_q = +227.7 patient-days, ΔS_r = +3.5 index-days, ΔS_u = −708 units of extraordinary capacity, the last being the difference between the 290 and 998 units of Table 2), ML-MPC is superior if and only if $\alpha/\gamma < 3.109 - 0.015\cdot(\beta/\gamma)$; at the adopted weights $\alpha/\gamma = 0.375$ against a threshold of 3.106, a margin of 8.3-fold. This frontier is exact for the two trajectories whose accumulated sums enter the calculation, that is, for policies held fixed while the weights vary. ML-MPC is not fixed in that sense: it minimizes J, so a change of weights changes its trajectory and hence its accumulated sums, and the frontier does not transfer automatically to the re-optimized policy. A separate numerical sweep was therefore run with ML-MPC re-optimized in each cell. ML-MPC attained the lowest objective-function value in 14 of the 20 cells. The robust policy prevailed in the three cells where $\alpha/\gamma$ equals 5 or 10, that is, above the threshold predicted by the fixed-trajectory frontier. In the remaining three cells, all with $\alpha/\gamma \leq 0.19$, the naive fixed-capacity policy prevailed instead: there the cost of extraordinary capacity so dominates the cost of queueing that ML-MPC stops allocating within its four-day horizon, a degeneracy of the receding-horizon formulation that the weight sweep exposed. The agreement with the predicted threshold is therefore an empirical finding on the grid examined, confined to the region where the controller remains active, and not a demonstration that the closed-form condition governs the re-optimized controller.

### 3.4 Uncertainty of the Differences

Resampling the demand trajectory in moving blocks and re-executing all six policies on each replicate gave 95% intervals for four quantities: the difference in objective-function value and the difference in total extraordinary capacity, each for ML-MPC against the robust policy and for ML-MPC against the naive fixed-capacity policy. All four excluded zero at all three block lengths. No other pairwise comparison was tested. Against the robust policy, the median reduction in the objective function was 58.3% (95% CI 53.3 to 63.1) for blocks of 14 days, 58.1% (53.9 to 62.3) for 30 days and 57.6% (54.1 to 60.8) for 60 days; the corresponding differences in extraordinary capacity were −705 (−769 to −642), −702 (−778 to −630) and −700 (−786 to −605) units. Against the naive fixed-capacity policy the median reductions were 45.9% (40.4 to 51.5), 45.7% (41.0 to 50.5) and 45.1% (41.3 to 49.6). The medians vary by less than one percentage point across block lengths, and the intervals widen as blocks lengthen, as expected when fewer independent blocks are drawn. These intervals are conditional on the simulation model and on the resampling mechanism itself, which reorders pairs of arrivals and forecasts and so creates mismatches at block junctions. They quantify variability under that mechanism and must not be read as prospective validation.

**Figure 5.** Trajectories of queue length (top), accumulated risk index (middle) and extraordinary capacity (bottom) for the six policies over the test period (2021–2022).




**Source:** Prepared by the authors (2026).

### 3.5 Start of the Modeling Window

Repeating the entire protocol with training beginning in 2010 instead of 2015, holding hyperparameters, validation year, test period and policies fixed, left the conclusion unchanged. Forecast accuracy was essentially the same (test MAE 0.9956 against 0.9941), and ML-MPC remained the lowest-cost policy in both protocols with a practically unchanged objective-function value (3,526.7 against 3,527.0). The ordering of policies was preserved. One policy did change materially: trained on 2010–2019, the historical-mean rule allocated no extraordinary capacity at all and collapsed onto the baseline, because a mean taken over a decade of rising demand never exceeds a base capacity of 1.597. This illustrates a limitation of rules built on historical averages that do not accommodate trend, and it is the reason such a rule is reported here as a comparator rather than as a candidate policy.

### 3.6 Base Capacity under the Test-Period Length of Stay

Mean length of stay fell from 11.7 days in the training period to 8.40 days in the test period (Section 4). Holding the 19-patient census constant, the test-period mean length of stay of 8.4028 days would imply a base capacity of b = 19/8.4028 ≈ 2.261 discharges per day and a utilization of 0.66 instead of 0.94. This value lies above the range covered in Section 3.3, so the scenario was run separately, with the recursive forecasts, hyperparameters and policy definitions unchanged. ML-MPC again attained the lowest objective-function value (1,260 against 3,101 for the robust policy and 6,072 for the naive fixed-capacity policy), using 38 units of extraordinary capacity against 289 and 727. Its advantage widened rather than narrowed: 59.4% below the robust policy, against 58.5% at the calibrated capacity, and 79.3% below the naive policy, against 46.0%. The ordering of the remaining policies, however, reverses. With utilization at 0.66 there is little

queue to relieve, so allocating extraordinary capacity becomes largely waste: the naive fixed-capacity policy moves from second best to worst, while the baseline moves from worst to second best. The historical-mean and point-forecast policies allocate nothing at all and coincide exactly with the baseline, since their forecasts never exceed 2.261. Two observations follow, both restricted to the scenarios examined. ML-MPC attained the lowest objective-function value in every configuration tested here, including this one at a utilization of 0.66, so its advantage is not confined to the calibrated capacity; this does not establish independence from calibration in general, nor superiority at utilizations outside the range examined. The relative merit of the simpler policies, by contrast, is entirely contingent on that calibration, and no ordering among them should be read as a property of the policies themselves.

## 4 DISCUSSION

This study compared six capacity allocation policies for inpatient stroke care in a simulation calibrated from real admission records. Across the 727-day test period, ML-MPC attained the lowest value of the objective function while using less than a third of the extraordinary capacity consumed by the robust policy, and the ordering of policies survived resampling of the demand trajectory, variation of the risk parameters and of the objective weights, and a change in the start of the modeling window.

The comparison of the six policies locates where the operational gain arises. A naive rule that allocates one extra unit every day, with no forecast, already brings the simulated queue down by an order of magnitude, so part of the achievable improvement follows simply from having extraordinary capacity available. The two point-forecast policies, by contrast, allocate on very few days and behave close to the baseline, because their forecasts seldom exceed base capacity. Forecast accuracy alone therefore produced little operational benefit here, consistent with the narrow margin by which the Random Forest improved on a historical-mean benchmark. What distinguishes ML-MPC is that the forecast, a safety margin and the capacity constraint enter a sequential optimization together; the design does not separate the causal contribution of these components, and no claim is made that any one of them accounts for the result on its own.

The robust and ML-MPC policies trade off differently rather than one dominating the other. The robust policy provisions against a high quantile of the historical distribution and so keeps the queue lower through the body of its distribution, at roughly three times the extraordinary capacity. ML-MPC concentrates allocation on the days its forecast flags and withholds it otherwise, accepting a larger routine queue for a much smaller resource footprint. Which profile is preferable is a management decision this study does not make; Section 3.3 delimits the region of that trade-off in which the reported ordering holds.

Several limitations bound these findings. The study draws on a single hospital, so the demand process, the case mix and the capacity constraints are those of one service in one region, and external validity is unknown. No queue for an inpatient bed was observed at the center: the queue, the wait proxy and the risk index are constructs of the model, and the risk index in particular is a relative mathematical quantity that was not validated against any clinical outcome and does not measure safety, disability or mortality. A further limitation concerns the comparability of the test period. The service that generated the 2021–2022 test set is not the one that generated the 2015–2019 training set: mean length of stay fell from 11.7 to 8.4 days, in-hospital mortality from 21.0% in 2019 to 12.7% in 2022, and the share of ischemic stroke rose from 76.8% to 84.2%. The difference in length of stay begins in 2020 and persists, although these data do not identify its cause. It raises uncertainty about the representativeness of the fixed base-capacity estimate: holding the census constant, a shorter stay implies a higher service rate. Effective capacity in the test period was not independently observed, and the 19-patient census is itself the 90th percentile of a different interval, so the direction and magnitude of any resulting bias cannot be established from these records. Section 3.6 tests the consequences directly. Base capacity is an estimate derived from the observed census rather than an administrative figure, and the model holds it constant, which the 2020 data show is not innocuous: mean length of stay fell from 10.4 to 8.7 days that year, a reduction of 16.3%, which would raise b by 19.5% if the number of beds entering Little's Law were held constant, an assumption these data do not verify. Because demand grew by 5.9% per year over the modeling window, a capacity calibrated on 2015–2022 may lose representativeness in later years, although the present data do not establish that it has. The weights of the objective function were set by judgment and carry no monetary meaning, so differences in J are differences in a constructed criterion. Exact enumeration of the MPC subproblem does not scale to multiple resources or longer horizons, and the weight sweep exposed a limitation of the receding-horizon formulation: in three configurations where extraordinary capacity was heavily penalized relative to queueing, ML-MPC allocated no additional capacity, despite a fixed-capacity policy achieving a lower objective value over the full evaluation period. This behavior is consistent with horizon myopia, although the inequality $\gamma > 4\alpha$ alone does not establish the condition, because the objective also includes the accumulated operational penalty. Longer horizons or terminal-state penalties warrant investigation. Finally, no patient was exposed to any of the evaluated policies,

so every reported difference is an output of the model rather than a measure of effect.

Three directions follow. The first is external validation: applying the same protocol to other services, with different case mixes and capacity regimes, would show whether the ordering reported here is a property of the method or of this series. The second is prospective calibration, including a time-varying base capacity (the thrombolysis rate at the study center grew from 1.2% in 2006 to about 10% in 2022 [18] over the same period in which length of stay fell, although these data do not establish that one caused the other, since changes in case mix, admission and discharge criteria and service organization could each account for part of the difference; whichever the cause, a constant b cannot represent a length of stay that drifts), together with nonlinear cost functions and longer horizons with a terminal cost on the queue state. The third, and the precondition for any operational use, is evaluation with hospital managers and clinical staff: the criterion optimized here was constructed by the authors, and whether it reflects the trade-offs that a stroke service actually faces is a question that only the people running that service can answer.

## 5 CONCLUSIONS

This paper compared six capacity allocation policies for inpatient stroke care in a simulation calibrated from 6,224 real admissions at a public neurological reference hospital in Southeastern Brazil. Base capacity, mean length of stay and the demand series were estimated from the admission records. The risk parameters and the weights of the objective function were not: they are design choices made by the authors, and the sensitivity analyses delimit how far the results depend on them.

Over the test period (2021–2022) the simulated service operated at 94% utilization, and the non-adaptive baseline sustained a mean queue of 4.34 patients. ML-MPC attained the lowest value of the objective function, 58.5% below the robust policy while using 70.9% less extraordinary capacity, and 46.0% below a naive fixed-capacity policy with 60.1% less extraordinary capacity; the robust policy kept a lower queue in the body of its distribution, at roughly three times the extraordinary capacity. These differences persisted under resampling of the demand trajectory, variation of the risk and cost parameters, and a change in the start of the modeling window.

Within the conditions of the model, the results indicate that how a forecast, its uncertainty and the capacity constraints enter sequential allocation decisions matters more than forecast accuracy alone. They do not demonstrate improved management of a real service, nor maintained clinical safety, nor financial savings: no patient was exposed to any evaluated policy, the risk term is a relative index that was not clinically validated, and the objective-function weights were set by judgment and carry no monetary meaning. Prospective validation, external replication and evaluation with hospital managers and clinical staff are required before any operational use.

**Author contributions.** Guilherme Coelho: Writing – review & editing. Daniel Correa da Silva: Software. Maurício Wesley Perroud Júnior: Resources. Clarimar José Coelho: Conceptualization, Formal analysis, Methodology, Project administration, Validation, Visualization, Writing – review & editing.

**Acknowledgments.** The authors thank the Hospital Estadual Sumaré (HES) and the São Paulo State Health Secretariat for providing access to the anonymized administrative dataset. The authors also acknowledge the UNICAMP Research Data Repository (REDU) for data preservation and open-access deposition.

**Funding.** This study did not receive external funding. Guilherme Coelho is supported by a doctoral fellowship at the School of Medical Sciences (FCM), University of Campinas (UNICAMP). The funders had no role in study design, data collection, analysis, decision to publish, or preparation of the manuscript.

**Data and materials availability.** The analysis pipeline, the independent verification script that recomputes every reported quantity from the admission records, the sensitivity grids, the formal specification of the ML-MPC controller and the chart data underlying Figures 1 to 5 are deposited at Zenodo under doi: https://doi.org/10.5281/zenodo.23985198. The patient-level admission records are not redistributable; the conditions of access are those stated in Section 2.6.

**Conflicts of interest.** The authors declare no conflicts of interest.

**Use of artificial intelligence.** The authors used an AI assistant (Claude, Anthropic) in preparing this revision: drafting and language revision, implementation and debugging of the simulation and forecasting code, generation of figures, and preparation of the response to the reviewers. The tool did not define the research question, did not access clinical records beyond the de-identified series of Section 2.1, and did not originate the clinical interpretation. All outputs were reviewed by the authors, and every reported quantity was recomputed from the raw records by the independent verification script supplied with this submission. No AI tool is listed as an author; the authors take full responsibility for the content, the accuracy and the integrity of the manuscript.

## REFERENCES


1. Ministry of Health of Brazil. Hospital Information System of the SUS (SIH/SUS): admissions for stroke, 2018–2022 [Internet]. Brasília: DATASUS; 2023. [cited Apr 2026]. Available from: http://tabnet.datasus.gov.br
2. Kim J, Olaiya MT, De Silva DA, Norrving B, Bosch J, De Sousa DA, et al. Global stroke statistics 2023: availability of reperfusion services around the world. Int J Stroke. 2024;19(3):253-70. doi: https://doi.org/10.1177/17474930231210448
3. Feigin VL, Brainin M, Norrving B, Martins S, Sacco RL, Hacke W, et al. World Stroke Organization (WSO): Global Stroke Fact Sheet 2022. Int J Stroke. 2022;17(1):18-29. doi: https://doi.org/10.1177/17474930211065917
4. Saver JL. Time is brain: quantified. Stroke. 2006;37(1):263-6. doi: https://doi.org/10.1161/01.str.0000196957.55928.ab
5. Emberson J, Lees KR, Lyden P, Blackwell L, Albers G, Bluhmki E, et al. Effect of treatment delay, age, and stroke severity on the effects of intravenous thrombolysis with alteplase for acute ischaemic stroke: a meta-analysis of individual patient data from randomised trials. Lancet. 2014;384(9958):1929-35. doi: https://doi.org/10.1016/S0140-6736(14)60584-5
6. Stroke Unit Trialists' Collaboration. Organised inpatient (stroke unit) care for stroke. Cochrane Database Syst Rev. 2013:CD000197. doi: https://doi.org/10.1002/14651858.cd000197.pub3
7. Lotufo PA, Goulart AC, Passos VMA, Satake FM, Souza MFM, França EB, et al. Cerebrovascular disease in Brazil from 1990 to 2015: Global Burden of Disease 2015. Rev Bras Epidemiol. 2017;20(Suppl 1):129-41. doi: https://doi.org/10.1590/1980-5497201700050011
8. Green L. Queueing analysis in healthcare. In: Hall RW, editor. Patient flow: reducing delay in healthcare delivery. New York: Springer; 2006. p. 281-307. doi: https://doi.org/10.1007/978-0-387-33636-7_10
9. Little JDC. A proof for the queuing formula: L = λW. Oper Res. 1961;9(3):383-7. doi: https://doi.org/10.1287/opre.9.3.383
10. Sudarshan VK, Brabrand M, Range TM, Wiil UK. Performance evaluation of emergency department patient arrivals forecasting models by including meteorological and calendar information: a comparative study. Comput Biol Med. 2021;135:104541. doi: https://doi.org/10.1016/j.compbiomed.2021.104541
11. Porto BM, Fogliatto FS. Enhanced forecasting of emergency department patient arrivals using feature engineering approach and machine learning. BMC Med Inform Decis Mak. 2024;24(1):377. doi: https://doi.org/10.1186/s12911-024-02788-6
12. 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. doi: https://doi.org/10.1186/s12911-022-01787-9
13. Hillier FS, Lieberman GJ. Introduction to operations research. 7th ed. New York: McGraw-Hill; 2001.
14. Ben-Tal A, El Ghaoui L, Nemirovski A. Robust optimization. Princeton: Princeton University Press; 2009. doi: https://doi.org/10.1515/9781400831050
15. Zhu T, Liao P, Luo L, Ye HQ. Data-driven models for capacity allocation of inpatient beds in a Chinese public hospital. Comput Math Methods Med. 2020;2020:8740457. doi: https://doi.org/10.1155/2020/8740457
16. Camacho EF, Bordons C. Model predictive control. 2nd ed. London: Springer; 2007. doi: https://doi.org/10.1007/978-0-85729-398-5
17. Doyle FJ 3rd, Huyett LM, Lee JB, Zisser HC, Dassau E. Closed-loop artificial pancreas systems: engineering the algorithms. Diabetes Care. 2014;37(5):1191-7. doi: https://doi.org/10.2337/dc13-2108
18. Coelho G, Gonçalves AL, Coelho CJ, Perroud Junior MW. Trends in hospitalization and mortality among stroke patients in a Brazilian public hospital: a 16-year retrospective analysis. Rev Bras Neurol. 2025;61(2). doi: https://doi.org/10.46979/rbn.v61i2.67441
19. Markus HS, Brainin M. COVID-19 and stroke: a global World Stroke Organization perspective. Int J Stroke. 2020;15(4):361-4. doi: https://doi.org/10.1177/1747493020923472
20. Breiman L. Random forests. Mach Learn. 2001;45(1):5-32. doi: https://doi.org/10.1023/a:1010933404324
21. Powell WB. Reinforcement learning and stochastic optimization: a unified framework for sequential decisions. Hoboken: John Wiley & Sons; 2022. doi: https://doi.org/10.1002/9781119815068
22. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Med. 2015;12(10):e1001885. doi: https://doi.org/10.1371/journal.pmed.1001885
23. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385:e078378. doi: https://doi.org/10.1136/bmj-2023-078378


## EDITORIAL NOTES


The Journal of Health Informatics is a scientific journal maintained by the Brazilian Society of Health Informatics (Sociedade Brasileira de Informática em Saúde), a non-profit scientific association dedicated to the development of Health Informatics and Digital Health in Brazil. The editorial process is conducted with scientific independence, free from institutional, commercial, or sponsor interference, and follows international best practices in publication ethics.

This article was double-blind peer reviewed by specialists in the field. The editorial team used artificial intelligence tools, under human supervision, in the initial screening, in checking the text, references and numbers and, when needed, in redrawing figures, charts and tables, which the authors approved in the proof. The tools and safeguards are described in the journal's Policy on the use of artificial intelligence. Editorial decisions were made by people.

SBIS Sociedade Brasileira de Informática em Saúde