Timing it Right: Balancing Inpatient Congestion versus Readmission Risk at Discharge

Pengyi Shi  
Krannert School of Management, Purdue University, West Lafayette, IN 47907, shi178@purdue.edu  
Jonathan E Helm  
Kelley School of Business, Indiana University, Bloomington, IN 47401 helmj@indiana.edu  
Jivan Deglise-Hawkinson, Julian Pan  
Lean Care Solutions Corporation Pte. Ltd., Singapore 139959 jivan, jp@leancaresolutions.com

When to discharge a patient plays an important role in hospital patient flow management as well as quality of care and patient outcomes. In this work, we develop a practical decision support tool to aid hospitals in managing the delicate balance between readmission risk at discharge and ward congestion. Our framework integrates a new prediction model that updates the readmission risk over a patient’s hospital stay with a large-scale Markov Decision Process (MDP) to optimize state-dependent discharge decisions. In the prediction model, we adapt and integrate several statistical methods to overcome three practical challenges. In the MDP, we overcome the curse of dimensionality by identifying useful structural properties that allow us to transform the MDP into a univariate optimization, leading to an efficient dynamic heuristic.

Through extensive counterfactual analyses, we demonstrate the value of our recommended discharge policy over our partner hospital’s historical discharge behavior. We then generalize the insights by applying the tool to a broad range of hospital types through a simulation study, and we identify two important operating regimes that explains when our dynamic discharge policy is most beneficial. Lastly, we discuss the implementation efforts of this discharge optimization tool at our partner hospital.

Key words: Readmission Prediction, Patient Flow Management, State-dependent Inpatient Discharge, Large-scale MDP, Approximation Algorithms, Tool Implementation

1. Introduction

A hospitalist makes many decisions that influence the cost of an inpatient stay...but probably none has more impact than “Should this patient go home today or tomorrow?”

–Cover story for American College of Physicians (ACP) Hospitalist, October 2014 (Colwell 2014).

This article raises the key tradeoff in making discharge decisions: “Under the Affordable Care Act, it is still in hospitals’ financial interest to discharge patients as soon as possible, but also to facilitate post-discharge care and prevent 30-day readmissions. Rather than just lowering LOS (length-of-stay), hospitals now aim to optimize it at the intersection of quality and cost.” How to balance this tradeoff has broad implications for patient flow, inpatient unit congestion, quality of care, and post-discharge risk, impacting all care providers from small community hospitals to major teaching hospitals.

Frequent overloading of inpatient units contributes to emergency department (ED) overcrowding (Proudlove et al. 2003), denial of ICU admission (Kim et al. 2016), cancellation of elective surgeries (Helm et al. 2011), and higher risk of mortality (Kuntz et al. 2014), among other consequences.
On the one hand, when inpatient units become congested, doctors frequently discharge existing patients early (Berry Jaeker and Tucker 2016, Kc and Terwiesch 2012, 2017). This practice alleviates overcrowding in the ward by shifting the burden to the early discharge patients, who may experience increased risk of readmission, mortality, and other adverse outcomes (Anderson et al. 2012, Kc and Terwiesch 2009, 2012). On the other hand, when occupancy levels are low, hospitals may keep patients longer (Anderson et al. 2011), which can have a positive impact on patient outcomes (Bartel et al. 2014). The balancing act between individual discharge risk and ward congestion has grown into a major stress point in the face of recent pressures to reduce readmissions while limiting LOS, as indicated in the article referenced above. Hospitals manage this tradeoff through ad-hoc practices that lack analytical decision support. OM literature has made significant strides in this area, though new research is needed to support the development of practical tools that can be implemented as part of a hospital’s workflow.

In this article, we develop a practical decision support tool to address two key research challenges: discharge risk prediction and dynamic discharge decision making. We test and implement this tool in close collaboration with a partner hospital. We focus on readmission risk since it is the target area of improvement in our partner hospital, though our framework is adaptable to other types of discharge risk such as mortality and adverse events. Figure 1 illustrates a conceptual framework for our decision support tool. We start from developing a prediction model that leverages several methods to overcome three main barriers: (1) excess zero count (most patients are not readmitted), (2) LOS correlates with both patient severity and readmission risk, (3) lack of heterogeneity in the commonly used Cox proportional hazard model; to be detailed in Section 1.1. We then integrate this prediction model into a discharge optimization framework that takes into account both ward congestion and individual discharge risk. This optimization forms a large-scale Markov Decision Process (MDP) that is intractable except in very small problem instances. We overcome the curse of dimensionality using a series of methods to reduce the impact of large action space and state space, leading to new and accurate heuristics. The final product of this work is a discharge decision tool that is in the process of a pilot implementation at a partner hospital.

Figure 1  Conceptual diagram of decision support development for discharge management.

1.1. The Industry Story
In this section, we discuss our partnership with hospital XYZ, during which we illustrate the widespread challenges faced by hospitals in balancing inpatient congestion and readmission risk. Hospital XYZ had contracted with a data analytics company, Lean Care Solutions (LCS), to help reduce
their readmissions. As a first step, LCS developed a tool that predicts each patient’s readmission risk by day after discharge. This tool helped reduce readmission rates by allowing Hospital XYZ to better target their limited resources to follow up with the highest risk patients at the appropriate time (e.g., when patients are at the greatest risk of readmission). With initial successes from the post-discharge tool, Hospital XYZ expressed to LCS a desire to exercise a second powerful lever in the readmission battle: pre-discharge interventions with a focus on discharge timing. Currently, early/late discharge is done in an ad-hoc manner without analytical decision support.

**Prediction.** In the exploratory phase, LCS found their prediction methods based on the classical Cox proportional hazard model insufficient to support discharge decisions. One obvious reason LCS discovered was that LOS is treated as an exogenous variable, leading to the misleading conclusion that extending length of stay for an individual patient results in higher readmission risk, since patients who stay longer tend to be sicker. We began a collaboration with LCS to address this endogeneity problem, and in the process discovered two other key challenges. First, while analyzing LCS’s method, we found that the prediction curves were being heavily distorted by the fact that many patients are never readmitted in the data, which conflicts with the underlying assumption of the basic Cox model that all patients will eventually be readmitted. Second, the Cox model produces a time to readmission curve that has the same baseline hazard function for all patients: the post-discharge readmission risk peaks at the same time for all patients. However, our data indicate some patients exhibit “early readmitter” behavior whereas others exhibit “late readmitter” behavior (see Figure 2 from our partner hospital). The ability to capture this heterogeneity is important because (1) having more detailed knowledge of when a patient is at risk enables the hospital to better target post-discharge follow-ups; (2) the medical literature has shown that the timing of readmission is correlated with readmission intensity (e.g., resource usage, burden on staff, Skolarus et al. (2015)). Further, readmission timing forms an important component of the patient flow model in our discharge optimization.

Discharge Decision. Even with a pre-discharge risk prediction tool, discharge planners still face complex decisions on how many patients and who to discharge on a given day. Discharge decisions must not only account for the risk of each patient, but also for each patient’s risk evolution over future days.
days in conjunction with current and future occupancy levels. The inpatient arrival day-of-week phenomenon further complicates discharge decisions. We explain this to our industry partners using the following simple illustrative example. Consider a patient who has a relatively high risk currently, but this risk is unlikely to improve significantly by keeping the patient longer. Then the best decision may be to discharge the patient now. The reverse may be true for a patient with lower discharge risk that may improve significantly by staying one day longer. This contradicts the industry belief, per our conversations with several hospitals, that a simple risk threshold is sufficient (i.e. discharge all patients when their risk drops below a certain level). In addition, the decisions are modulated by considering current and future occupancy levels rather than risk alone. These complexities necessitate a forward-looking, dynamic approach that cannot be easily intuited.

1.2. Overview and Main Contributions.
This paper contains both technical and practical contributions to the literature.

Technical Contributions.
- Prediction algorithm. In Section 3, we develop a method that dynamically predicts the trajectory of readmission risk during the patients stay in the hospital; i.e., how risk will evolve over time if the patient is discharged on the current day or any future day. To overcome the three aforementioned empirical barriers, we construct a Cox-cure model to correct for the excess zero count, develop a clustering and Expectation Maximization algorithm to capture patient heterogeneity, and adopt an instrumental variable (IV) approach to correct for LOS endogeneity. The prediction algorithm is validated on our partner hospital’s dataset with satisfactory performance.

- Optimization framework. In Section 4, we build a large-scale Markov Decision Process (MDP) that integrates the risk prediction and a patient flow model with re-entries. This MDP accounts for heterogeneity in patient class as well as how long each patient has currently spent in the hospital, which leads to a high-dimensional state and action space, differing from conventional models that assume memoryless service time distributions.

- Analytical results and heuristic. To overcome the curse of dimensionality, in Section 5 we prove structural properties of the MDP showing that discharge decisions should depend on marginal risk among all future days, contradicting the belief in the medical community that current absolute risk should be the criterion for discharge. We develop both a strong dominance and a more general weak dominance criterion to rank patients for discharge. In Section 6, we leverage a special case of the MDP which can be solved as a linear quadratic stochastic control problem. Combining novelly the closed form solutions for the optimal decision and value function of this special MDP with the patient ranking, we transform the original MDP into a heuristic univariate optimization that significantly reduces the computational complexity.

Practical contributions.
- Insights from case study. To gain buy-in for implementation, we start with a trace-based counterfactual in Section 7.1 that demonstrates Pareto dominance of our dynamic heuristic policy over
the historical discharge policy in occupancy and readmissions. Our policy properly identifies over 50% of patients that should have received an intervention (i.e. extended LOS) but did not. We also uncover an interesting secondary benefit – occupancy smoothing – exhibited only by the dynamic policy. We generalize these results to other hospitals through simulation analyses in Section 7.2, showing that our heuristic gains more benefit in hospitals/inpatient units (1) that are smaller, less congested, (2) where patients recover faster (e.g. elective surgery), and (3) exhibit greater day-of-week arrival variability. We explain these insights by uncovering two operating regimes: occupancy-driven versus quality-driven discharge regimes.

• Implementation. We worked with our partner hospital to develop, test, and implement a cloud-based tool that hospital users can receive discharge decision support; see Figure 3a for a snapshot of the main portal. The tool, which has recently been integrated into the hospital’s IT infrastructure and provider workflow, displays for each patient (1) discharge risk curve for future possible LOS, (2) recommended discharge date, and (3) post-discharge readmission timing risk curve; see Figure 3b for features (1)-(2), enlarged in Appendix F. We discuss implementation efforts and supporting evidence of the benefit from our tool in Section 8.

Figure 3 Screenshot of the discharge decision support web portal implemented in our partner hospital.

Remark 1. In practice, the discharge decisions are complicated and involve a variety of factors. We emphasize that our tool is meant to provide analytical support for discharge decisions. Doctors can still use their own discretion in discharging patients according to the individual patient’s condition and needs. Our tool is flexible to accommodate such deviations and can update the recommendations after incorporating the actual decisions; see details in Section 8.

2. Literature Review
We review three streams of literature relevant to our paper.

Empirical evidence on discharge and patient outcomes. A rich body of empirical research has provided evidence that hospitals tend to use discharge decisions to reduce inpatient unit congestion. Sharma et al. (2008) show that when hospitals become capacity constrained, patients are discharged early. Kc and Terwiesch (2017) and Berry Jaeker and Tucker (2016) also find indications
of early discharge when inpatient occupancy increases. In the ICU, Kc and Terwiesch (2012) find a similar “speedup” phenomenon during congested periods. Long and Mathews (2017) show that ICU occupancy impacts the less essential “boarding time,” but not the medically necessary LOS.

While early discharge can alleviate congestion and increase chance of admission for future patients, it compromises patient outcomes. Using a large dataset on CHF patients, Oh et al. (2017) find inpatient stays that are shorter than the CMS suggested LOS, are likely to exhibit a 1.1% greater risk of readmission. Kc and Terwiesch (2012), and Kc and Terwiesch (2009) find that patients discharged early exhibit increased risk of readmission, mortality, and other adverse outcomes. The medical literature has discovered similar findings between LOS and patient outcomes, e.g., Kuo and Goodwin (2011), Heggestad (2002), to name a few. On the other hand, Anderson et al. (2011) hypothesize that there may also be a phenomenon of keeping patients longer when occupancies are lower. Bartel et al. (2014) shows that keeping a patient one extra day can reduce mortality risk by nearly 6%. Oh et al. (2017) and Carey (2015) suggest that keeping patients longer can significantly reduce hospital costs. Our paper is largely motivated by these empirical studies, though it is important to differentiate our work from the these econometric studies, as our focus in Section 3 is on dynamic, personalized prediction of readmission risk evolution over a patient’s hospital stay. In addition, our paper develops an integrated tool that provides analytical support for balancing the tradeoff between ward congestion (inpatient LOS) and readmission risk.

**State-dependent discharge optimization.** Our modeling and discharge decision analysis connects with the literature on optimal service rate control. Within this area, several papers specifically study discharges in the hospital. Berk and Moinzadeh (1998) is one of the earliest papers to study the tradeoff between discharge risk and inpatient occupancy. The authors model patient care for a homogeneous population in two stages, where stage 1 is “not dischargable,” and stage 2 is a less critical stage in which early discharge can be exercised. They perform steady-state analysis under two fixed policies (with and without early discharge) and find that early discharge can reduce occupancy without sacrificing quality of care. In contrast to this work, our paper focuses on decision support, leveraging both prediction and control theory to determine optimal discharge actions for multiple patient classes. Motivated by their model, we also consider an extension to incorporate the two-stage health states in Appendix E.1. Crawford et al. (2014) develop a simulation study on the impact of inpatient discharge policies on ED congestion and readmission, where they evaluate the performance of three fixed discharge policies.

Chan et al. (2012) model the discharge decision in the ICU as a continuous time MDP so that, when a new patient arrives to a full ICU, doctors must decide which patient to discharge to free a bed. Our paper considers a discrete time model, which is more natural for the inpatient discharge setting, since discharges are usually processed once a day during rounds. Hence, we determine both which patients and how many patients to discharge, considering patient risk trajectories and current and future occupancy. We also track how long a patient has stayed, linking this to a LOS-dependent
risk. Chan et al. (2012) suggest developing a risk prediction model as an important area for future work, which is a major component of our paper. Ouyang et al. (2015) consider the joint decision of ICU admission and discharge decisions. The decision maker decides whether to admit an arriving patient to the ICU or to the general ward and also who to to discharge early if a patient needs to be admitted to a full ICU. An important insight the authors find is that the optimal decisions not only depend on the expected ICU benefit of a patient but also how long they will stay to get this benefit. This is similar to our finding that the discharge “desirability” of a patient depends on the magnitude of risk reduction in the futures days, not just the absolute risk level.

In the more general service operations literature, George and Harrison (2001), Ata and Shneorson (2006), Bekker and Boxma (2007) study optimal control of single-server queues with state-dependent service rates. Huang and Gurvich (2018) and Braverman et al. (2018) develop new frameworks to identify asymptotic optimal control in single-server queues with abandonment. Most of these service-rate works (i) provide an optimal rate, but not which customer(s) to discharge; and (ii) do not consider returns to service (readmission), except Chan et al. (2014), who use a fluid model to determine whether a speedup service rate should be used under different system states. Our paper considers a richer patient profile and readmission risk trajectory, and our discharge decisions explicitly account for the current LOS of each patient. In addition, Chan et al. (2014) focus on equilibrium analysis, while our focus is in on the optimal control.

Integrating prediction with optimization. There is a growing literature on integrating prediction models with operational optimization, i.e., data-driven optimization. For example, Barnes et al. (2015) develop a tree-based prediction model on a patient’s likelihood of discharge by 2pm or midnight each day, which then can be used to rank patients for task prioritization. Hu et al. (2018) propose using a new patient risk score to initiate proactive ICU transfer, and demonstrate the benefits via a simulation study. In the retail sector, Ferreira et al. (2015) develop a demand prediction model for an online retailer, that is used as input into a price optimization model to maximize revenue; our paper has a similar spirit. Importantly, Ferreira et al. (2015) implement this algorithm with their partner and conduct a field experiment.

3. Dynamic prediction of patient discharge risk

In this section, we develop a new model to predict time to readmission that extends the classical Cox proportional hazard model (Cox 1992). Our approach, which addresses the three empirical challenges, is summarized in Figure 4. In Stage 1, we use a cure model to address excess zero count. This stage predicts the overall probability that a patient is eventually going to be cured or readmitted. For patients who are not cured (will be readmitted), we move to Stage 2, where we use a mixture model to capture patient heterogeneity with clustering and estimate different time to readmission curves for each cluster. We use an instrumental variable (IV) method, replacing LOS with a predicted LOS in Stage 1 and Stage 2, as a pre-processing stage to correct for endogeneity.
In Section 3.1, we give an overview of our Cox-cure mixture model with IV correction. In Section 3.2, we introduce an expectation maximization (EM) algorithm to estimate parameters for the Cox-cure mixture model. In Section 3.3 we specify the IV method. In Section 3.4 we provide a detailed parameterization and validation of our prediction model using data from our partner hospital. A brief review of the basic Cox model is available in Appendix A.1.

3.1. Cox-cure model with mixture component

In this section we develop a cure model framework (see, for example, Yu (2008), Bardhan et al. (2014)) to address the issue that a large proportion of patients are never readmitted, which can heavily distort the estimation of the Cox model. The cure model is a two-stage approach: the first stage predicts whether each patient is cured or not; if the patient is not cured, we model the time to readmission using a Cox model with a mixture component to capture patient heterogeneity in the second stage. Let \( C_i \) denote patient \( i \)’s cure status, where

\[
C_i = \begin{cases} 
1, & \text{cured;} \\
0, & \text{uncured.} 
\end{cases}
\]

Conditioning on \( C_i = 0 \), a patient belongs to one of \( W \) clusters with probability \( \pi_w \) for \( w = 1, \ldots, W \); \( \{\pi_w\} \)'s are population membership probabilities with \( \sum_{w=1}^{W} \pi_w = 1 \). Each cluster \( w \) has an associated time-to-readmission proportional hazard rate function \( h^w(t; i) \). For each patient \( i \), we introduce the individual membership variable, \( Z_i \), to denote which cluster patient \( i \) belongs to, with \( Z_i \) being drawn from a multinomial distribution with mixing weights \( \{\pi_w\} \).

Both the cured status \( C_i \) and the membership variable \( Z_i \) are latent variables; i.e. unobservable. To perform parameter estimation with these latent variables, we develop an EM framework detailed in Section 3.2. Next, we specify the parametric models for Stages 1 and 2.

**Parametric models.** In the first stage, we assume that the probability of a patient being cured follows a logit model. For patient \( i \) with features \( Y_i \), her cure probability \( \theta_{0i} = \mathbb{P}(C_i = 1) \) follows

\[
\log \left( \frac{\theta_{0i}}{1 - \theta_{0i}} \right) = Y_i \xi,
\]

where \( \xi \) is the set of coefficients associated with the individual patient’s characteristics and risk factors, denoted as \( Y_i \):

\[
Y_i = (\log(LOS_i), Y_i^e).
\]
Here, \( Y_{i}^c = (Y_{i,1}^c, \ldots, Y_{i,K-1}^c) \) denotes \( K - 1 \) exogenous variables such as patient age, gender, and medical specialties. During the model training phase (for parameter estimation), we replace the actual \( \log(LOS) \) observed from data with \( \log(LOS_i) \) predicted from a linear regression in Stage 0.

For the second stage, we use the Cox model. For patient \( i \) in cluster \( w \), the hazard rate is

\[
h^w(t; i) = h^w_0(t) e^{Y_i \beta^w}, \quad w = 1, \ldots, W. \tag{2}
\]

The baseline hazard rate function \( h^w_0(t) \) follows the Weibull form (see Equation (28) in Appendix A.1) with cluster-dependent parameters \( \lambda^w \) and \( k^w \). The coefficients \( \beta^w = \{\beta^w_1, \ldots, \beta^w_K\} \) are also cluster-dependent. \( Y_i \) is the same as the first stage, but \( \beta^w \) can be different from \( \xi \).

**Readmission probability prediction.** For a new patient \( i' \) not used for model training, we predict her readmission timing by first obtaining the cure probability \( \hat{\theta}_{0w} \) from the logit model (1). Then, the \( t^* \)-day cumulative readmission probability for this patient \( i' \) equals \( (1 - \hat{\theta}_{0w}) H(t^*; i') \), where

\[
H(t; i') = \sum_{w=1}^{W} \pi_w \left( 1 - \exp \left( - \int_{0}^{t} h^w(u; i') du \right) \right). \tag{3}
\]

The probability of being readmitted on a particular day \( t \) equals \( (1 - \hat{\theta}_{0w})(H(t; i') - H(t - 1; i')) \).

### 3.2. Parameter estimation for the mixture Cox-cure model

In this section, we specify the likelihood functions and provide a high-level description of the EM algorithm for parameter estimation; see full details in Appendix A.2. Let \( J_i \) denote the number of readmission events for patient \( i \). The complete-data likelihood function is given by

\[
L = \prod_{i} \left( \theta^1_{0i} \mathbb{1}\{C_i=1\} \cdot (1 - \theta_{0i}) \mathbb{1}\{C_i=0\} \right) \cdot \prod_{i \in \{J_i=0\}} \prod_{w=1}^{W} \left( \pi_w \cdot \exp \left( - \int_{0}^{T_i} h^w(u; i) du \right) \right) \mathbb{1}\{C_i=0, Z_i=w\} \\
\quad \cdot \prod_{i \in \{J_i \neq 0\}} \prod_{w=1}^{W} \left( \pi_w \cdot \left( H^w(t_i; i) - H^w(t_i - 1; i) \right) \right) \mathbb{1}\{C_i=0, Z_i=w\}, \tag{4}
\]

where \( H^w(t; i) = 1 - \exp \left( - \int_{0}^{T_i} h^w(u; i) du \right) \), and \( T_i \) is the data censoring time minus the patient’s discharge time. The first product corresponds to the logit (cure) model, and the second and third products correspond to the mixture model. For the second product, a patient \( i \) without a readmission event \( (J_i = 0) \) could be uncured but the readmission event is “censored” at \( T_i \) (not observed in the data set). If the patient belongs to cluster \( w \), i.e., \( \{C_i = 0, Z_i = w\} \), the censoring probability equals \( \exp \left( - \int_{0}^{T_i} h^w(u; i) du \right) \). For the third product, a patient with a readmission event, \( (J_i \neq 0) \), is uncured. If she belongs to cluster \( w \), the probability of a readmission on day \( t_i \) is \( (H^w(t_i; i) - H^w(t_i - 1; i)) \).

The labels for the latent variables, \( \mathbb{1}\{C_i = 1\} \) or \( \mathbb{1}\{Z_i = w\} \) in (4) are unobservable. The cured status \( C_i \) is partially observable (Yu 2008). That is, if a patient is readmitted, we know she must have been uncured; however, if this patient is not readmitted, we do not know whether she is cured or she is uncured but the readmission time is censored. To address this, we develop an EM algorithm, which iteratively updates the estimation of parameters and distributions on the latent variables to
maximize the expected log likelihood function. In each iteration $n$, we define two surrogates for the cured probability and membership probability:

$$P_{0,i}^{(n)} \sim \mathbb{P}(1 \{C_i = 1\}), \quad T_{w,i}^{(n)} \sim \mathbb{P}(1 \{Z_i = w|C_i = 0\}), \ w = 1, \ldots, W;$$

$$P_{0,i}^{(n)} + (1 - P_{0,i}^{(n)}) \sum_{w=0}^{W} T_{w,i}^{(n)} = 1.$$ 

In the **E-step** of the EM algorithm, we update $P_{0,i}^{(n)}$ and $T_{w,i}^{(n)}$ using Bayesian estimation. Then, using the updated surrogates and the log transformation on (4), we approximate the expected log likelihood function as

$$E[LL] \approx \sum_i \left( P_{0,i}^{(n)} \log \theta_{0,i} + (1 - P_{0,i}^{(n)}) \log (1 - \theta_{0,i}) \right) + \sum_i (1 - P_{0,i}^{(n)}) \sum_{w=1}^{W} T_{w,i}^{(n)} \log \pi_w$$

$$+ \sum_i (1 - P_{0,i}^{(n)}) \sum_{w=1}^{W} T_{w,i}^{(n)} \left( -1 \{J_i = 0\} \int_0^{T_i} h^w(u; i) du + 1 \{J_i > 0\} \log \left( H^w(\tau_i + 1; i) - H^w(\tau_i; i) \right) \right). \quad (5)$$

Then, we update estimation of the parameters $\Psi = \{\xi, \{\pi_w\}, \{\lambda^w\}, \{k^w\}, \{\beta^w\}\}$ by maximizing the log-likelihood function (5) in the **M-step**. See details in Appendix A.2.

### 3.3. Correcting endogeneity of LOS with instrumental variables (IV)

Patient severity may be positively correlated with both LOS and the dependent variables $Y_i$ since sicker patients tend to stay longer and are also more likely to be readmitted. Thus, treating LOS as an exogenous variable can lead to the incorrect conclusion that longer LOS results in higher readmission risk. To address this issue, we follow the IV technique developed in Bartel et al. (2014), replacing log(LOS) with log($\hat{\text{LOS}}_i$) predicted by a linear regression on exogenous features $Y_i^e$ (which are included in the first- and second-stage models) and additional IVs, $IV_i$, which only appear in Stage 0. The linear regression on log(LOS) follows:

$$\log(\text{LOS}) = \phi Y_i^e + \zeta IV_i. \quad (6)$$

After obtaining the estimates for $\phi$, we replace the variable log(LOS) with the predicted value, log($\hat{\text{LOS}}_i$), in $Y_i$, and estimate the parameters in the Cox cure model.

For the IV method to be valid, the IVs must meet the following criteria: (i) relevance – log(LOS) depends on $IV_i$, and (ii) exclusion – $IV_i$ should be uncorrelated with the unobservables in the readmission risk. In their econometric study, Bartel et al. (2014) propose using admission day-of-week indicators as the IV in the regression (6). The rationale is that physicians prefer not to keeping patients over the weekend. Thus, a patient who would have stayed for 4 days in the hospital may get discharged early if she is admitted on Tuesday due to such operational considerations. In addition to the admission day-of-week indicators, we have tested other candidates such as discharge day-of-week and the average occupancy level on the day of discharge. In Section 3.4, we compare the Area Under Curve (AUC) performance of including different combinations of IVs and select the best performing variables in our final implementation. We conclude with two remarks.
Remark 2 (IV in Non-linear Models). The IV method is generally used in a linear regression framework (e.g. when Stage 0 and Stage 1 are both modeled by linear regressions). In our context we use this method in conjunction with the non-linear Cox survival model, where there is no theoretical guarantee except when the hazard rate function is linear in $Y_i$ (MacKenzie et al. 2014, Zheng et al. 2017). Despite this fact, IV methods have been promoted in survival analysis, with various numerical studies reporting success in correcting bias with IVs, e.g., Atiyat (2011), Tian (2016).

Remark 3 (Counterfactual). Our prediction analysis is connected with a rapidly growing research area using machine learning to estimate individual treatment effects (Athey 2017, Wang et al. 2017), where extending/shortening LOS is the treatment in our setting. Hartford et al. (2016) proposed a similar idea of using IVs and a two-stage process to perform counterfactual predictions, where both stages are modeled by deep neural networks. Our prediction model can also be considered as a two-stage process, where the second stage is a non-conventional model (the mixture Cox-cure model). The prediction power of our tool may be further improved if the Stage 0 linear model is replaced by a nonlinear model, say, neural nets; we leave this to future research.

3.4. Prediction model validation and implementation
We parameterize and validate the prediction model on a dataset from a partner hospital in the state of Indiana.

Data description. The data set from our partner hospital spans January 2010 to September 2017 with more than 200 available features, including patient demographics, psycho-social data, diagnostic information, post-discharge dispositions, indicators of secondary illnesses including depression and diabetes, and dates of admission and discharge for each visit (if multiple) for each patient. We exclude the binary predictors that are recorded for less than 5% of the population. We also exclude planned readmissions, expired patients, patients under the age of 18 (including newborns), and obstetrics and gynecology (OBGYN) patients. The final dataset included $n = 25601$ patients.

Model selection and validation. There are three key elements we use to fine tune the prediction model developed in Section 3.2: (1) the number of clusters, $W$, and cluster initialization in the EM algorithm, (2) exogenous features $Y^e$ (feature selection), and (3) instrumental variables (IV). Our performance metric for prediction capability is the Area Under the Curve (AUC). To choose the best model, we use the following algorithm and fine tune the three model components to identify the setting with the best AUC performance. We leave the details of the fine-tuning to Appendix A.3.

Algorithm for model parameterization and validation.
1. Data Partitioning: We employ a bootstrapping method to separate the data into training data (sampled from the original dataset) and testing data (the original dataset). We perform this random sampling 50 times, generating 50 different training datasets. Compared to cross validation typically used in the literature, bootstrapping allows for both a larger training and testing sample size, which is important in our setting since the number of readmitted patients is small compared to the total patient population (Austin and Steyerberg 2017).
2. **Model Estimation:** For each training dataset, we estimate \((\phi, \zeta)\) in Stage 0, and parameters
\[
\Psi = \{\xi, \{\pi_w\}, \{\lambda^w\}, \{k^w\}, \{\beta^w\}\}
\]
in Stages 1 and 2.

3. **Performance:** Using the model estimated in Step 2, we predict the \(t^*\)-day (e.g., \(t^* = 30\) or 90 days) cumulative readmission probability for each patient using Equation (3) for all patients in the testing data. We compare the predicted readmission probability with the actual readmission event to obtain the AUC score. In this step we use the actual LOS instead of the \(\log(\hat{L}\hat{O}\hat{S}_i)\) to make sure our discharge risk curve as a function of LOS is robust (predicts readmission well) to all possible LOS values, not just for the predicted value \(\log(\hat{L}\hat{O}\hat{S}_i)\). This provides a method to validate the risk prediction curve, not just a point estimate of LOS.

4. **Validation:** Repeat steps 2 through 3 for each of the 50 training sets, obtaining \(\Psi_i\) and \(AUC_i\) for training set \(i\). We then calculate an average parameter estimate, \(\overline{\Psi} = (1/50) \sum_{i=1}^{50} \Psi_i\), perform Step 3 again, and obtain \(AUC(\Psi)\). We fine tune the three elements mentioned above to find a model that performs well on both \(AUC(\Psi)\) and \(AUC = (1/50) \sum_{i=1}^{50} AUC_i\), as the latter is indicative of consistent performance of the model setting. We use \(\overline{\Psi}\) in the implementation.

**Summary of feature coefficients and estimated parameters.** The estimated coefficients \(\overline{\Psi}\) for selected main features are reported in Table 1. The signs of the coefficients for \(\log(LOS)\) suggest that extending LOS reduces the readmission risk; also see Figure 13 in Appendix A.4 for an illustration of the 90-day readmission risk as a function of LOS. For the 90-day readmission risk, \(AUC(\overline{\Psi})\) is 69.42%, and \(\overline{AUC}\) is 68.17% (±1.15% for the 95% confidence interval). For the 30-day readmission risk, \(AUC(\overline{\Psi})\) is 69.15%, and \(\overline{AUC}\) is 66.17% (±1.04%). This AUC performance is comparable with those reported in the literature. In Sections 7.1 and 8 we show numerical and empirical evidence that this prediction tool can help our partner hospital identify more patients at risk and has improved prediction metrics since the pilot implementation.

4. **Modeling framework for discharge decision optimization**
In this section, we formulate an infinite-horizon, average cost MDP for optimal discharge decisions based on the predicted risk trajectories from Section 3. Note that this discharge decision framework is designed for use with patients whose medically LOS falls into a normal range (e.g. 1–15 days), which compromises the majority of inpatients. We are not suggesting that our model should control the discharge for patients having excessively long LOS or complicated reasons for remaining in the hospital, whom should be handled on a case-by-case basis based on the doctor’s discretion.

4.1. **Patient flow model**
Figure 5 depicts the patient flow model. New patients arrive to a hospital ward with \(N\) beds according to a time-nonhomogeneous Poisson process with a periodic arrival rate function \(\lambda(t)\). At this point, we assume that the period is one day, with
\[
\Lambda = \int_0^1 \lambda(s)ds = \int_t^{t+1} \lambda(s)ds
\]
Features | Meaning | Stage 1 | Stage 2
--- | --- | --- | ---
\( \pi \) | Mixing weights | - | 0.53 0.47
\( \lambda \) | Weibull baseline hazard | - | 10.10 26.67
\( k \) | Weibull baseline hazard | - | 1.65 3.81
Cure prob | Proportional hazard | \( \xi \) | \( \beta_1 \) | \( \beta_2 \)

<table>
<thead>
<tr>
<th>Feature</th>
<th>Meaning</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \log(\text{LOS}) )</td>
<td></td>
<td>0.49 -0.24 -0.65</td>
</tr>
<tr>
<td>( \text{is_readmit} )</td>
<td>Visit as readmission</td>
<td>-0.53 0.05 0.10</td>
</tr>
<tr>
<td>( \text{Dis_Disposition1} )</td>
<td>Home health</td>
<td>-0.86 0.25 0.28</td>
</tr>
<tr>
<td>( \text{AttendSpclty4} )</td>
<td>Pulmonary Medicine</td>
<td>-0.37 -0.18 0.30</td>
</tr>
<tr>
<td>( \text{AttendSpclty5} )</td>
<td>Internal Medicine</td>
<td>-0.40 -0.23 0.36</td>
</tr>
<tr>
<td>( \text{CHF1} )</td>
<td>Congestive Heart Failure</td>
<td>-0.41 -0.03 0.11</td>
</tr>
<tr>
<td>( \text{ICD2} )</td>
<td>R00-R99</td>
<td>0.72 -0.14 -0.36</td>
</tr>
<tr>
<td>( \text{ICD3} )</td>
<td>S00-T98</td>
<td>0.57 -0.02 -0.21</td>
</tr>
<tr>
<td>( \text{ICD5} )</td>
<td>N00-N99</td>
<td>0.77 0.26 -0.31</td>
</tr>
<tr>
<td>( \text{ICD6} )</td>
<td>M00-M99</td>
<td>0.75 -0.05 -0.35</td>
</tr>
<tr>
<td>( \text{ICD8} )</td>
<td>J00-J99</td>
<td>0.38 -0.12 -0.15</td>
</tr>
</tbody>
</table>

Table 1  Estimated coefficients for \( \text{LOS} \) and selected main features. Note that \( \xi \) are for Stage 1 cure probability. Thus, a positive coefficient corresponds to a reduction in the readmission risk; vice versa. In contrast, \( \beta_1 \) and \( \beta_2 \) are for Stage 2 Cox model. Thus, a positive coefficient corresponds to an increase in the readmission risk.

denoting the exogenous daily arrival rate. We study the impact of the day-of-week arrival variability in Section 7.2. Here, each patient’s \( \text{LOS} \) depends on the discharge action and is no longer an exogenous variable as in conventional queueing models. Once a patient is discharged, she is either cured, or will be readmitted with probability that depends on their risk class and \( \text{LOS} \) at discharge.

![Patient flow model of the hospital ward.](image)

**Figure 5**  Patient flow model of the hospital ward.

We assume that each patient belongs to a discharge risk class \( m \) with probability \( p_m \), where there are \( M \) possible classes; i.e. \( m = 1, \ldots, M \), and \( \sum_{m=1}^{M} p_m = 1 \). Let \( r(m; j) \) be the probability of readmission for a class \( m \) patient with \( \text{LOS} \) of \( j \), which can be estimated from the first-stage cure model (1) in Section 3.1. Let \( q(t; m, j) \) denote the probability that this class \( m \) patient will be readmitted on day \( t \) post-discharge. We have

\[
q(t; m, j) = r(m, j) \cdot \left( H(t; m, j) - H(t - 1; m, j) \right), \quad \sum_{t=1}^{T} q(t; m, j) = r(m, j),
\]

where \( T \) denotes the maximum time we count a patient visit as a readmission (90 days in this paper), \( H(t; m, j) \) is the cumulative probability function estimated from the second-stage Cox model (2) and is appropriately normalized by setting \( H(T; m, j) = 1 \). Note that a class \( m \) patient may become a
different class $\tilde{m}$ upon readmission, with probability $\delta_{m,\tilde{m}}$, since prior readmission is found to be an important factor in our prediction model. We define $q_m(t; m, j)$ as

$$q_m(t; m, j) = \delta_{m,\tilde{m}} \cdot q(t; m, j), \quad \sum_{\tilde{m}} \delta_{m,\tilde{m}} = 1. \tag{9}$$

The $M$ risk classes can be obtained by grouping patients with similar predicted risk trajectories; these classes are not necessarily the same as the $W$ clusters from the Cox-mixture framework in Section 3, since the latter mainly captures the heterogeneity in the hazard function (readmission timing). We remove this risk classification and incorporate personalized patient risk in the numerical study and implementation in Sections 7 and 8.

4.2. Infinite Horizon Average Cost Problem

We formulate the discharge decision as a discrete-time, infinite-horizon average cost MDP. Below, we specify the system state, action, transition dynamics, cost and objective function.

**System state.** The system state is captured by the following $M \times (J + 1)$-dimensional vector:

$$X(t) = (X^0(t), X^1(t), \ldots, X^J(t))^\top,$$

where $X^j(t) = (X^{1,j}(t), X^{2,j}(t), \ldots, X^{M,j}(t))^\top$ and $X^{m,j}(t)$ denotes the number of class $m$ patients who have spent $j$ days in the system, $j = 0, \ldots, J$. Note that $X^{m,j}(t)$ includes both patients in service and waiting for a bed. We assume that a patient begins the treatment and recovery process immediately upon arrival, since patients still receive care even if not immediately placed in a bed.

**Action.** Each day the decision maker observes the system state at a decision epoch (e.g., at the time of rounds), and determines the number of patients to discharge. For notational convenience, we assume that this observation occurs at time 0 of each day. Mathematically, let $\{X(t), t \geq 0\}$ denote the system state, which is a continuous-time stochastic process. Let $X(k-)$ and $X(k)$ to denote the pre-action and post-action state at decision epoch $k$ (day $k$). Unless otherwise specified, we use $X_k = X(k-)$, $k = 0, 1, \ldots$ to denote the pre-action state. At decision epoch $k$, we take discharge action $D_k = (D^0_k, D^1_k, \ldots, D^J_k)^\top$, where $D^j_k = (D^{1,j}_k, D^{2,j}_k, \ldots, D^{M,j}_k)^\top$ and $D^{m,j}_k$ represents the number of discharges of class $m$ patients who have spent $j$ days in the system. A feasible discharge action satisfies $0 \leq D^{m,j}_k \leq X^{m,j}_k$, $\forall m, j$.

**Remark 4.** In our general formulation, we do not restrict the action space. i.e. all patients are potentially “dischargeable,” although the high readmission risk for short LOS patients would prevent their discharge in reasonable parameterizations. In Appendix E.1, we discuss an extension where a patient progresses through a critical stage (non-dischargeable) and then a stable stage as in Berk and Moinzadeh (1998). We show that our main theorem and structural properties still hold in this extended model. For our numerical study, however, we lack data to parameterize the two stages, and we focus on the formulation presented here.
Transition dynamics. Let $A_k^{m,0}$ denote the number of new arrivals belonging to class $m$ in period $k$ (between decision epochs $k$ and $k + 1$), and $A_k^{'m,0}$ denote the number of readmissions belonging to class $m$. For each $m$, the state evolution is:

$$X_{k+1}^{m,0} = A_k^{m,0} + A_k^{'m,0}; \quad (10)$$

$$X_{k+1}^{m,j} = X_k^{m,j-1} - D_k^{m,j-1}, \quad j = 1, \ldots, J. \quad (11)$$

Equation (10) captures the arrivals to the hospital ward in period $k$. Equation (11) says that patients who have stayed $j - 1$ days in period $k$ become patients who have stayed $j$ days in period $k + 1$, except for those who are discharged, $D_k^{m,j-1}$.

The total number of exogenous arrivals $\sum_{m=1}^{M} A_k^{m,0}$ follows a Poisson distribution with mean $\Lambda$. By Poisson splitting, $A_k^{m,0}$ is distributed as $\text{Poiss}(\Lambda_m)$ with $\Lambda_m = p_m \Lambda$. The readmission arrival stream, $A_k^{'m,0}$, on the other hand, depends on past discharge actions, i.e.,

$$A_k^{'m,0} = \sum_{t=1}^{T} \sum_{j=1}^{J} \sum_{m=1}^{M} \text{Bin}(D_{k-t}^{m,j}, q_m(t; \tilde{m}, j)),$$

where $\text{Bin}(\cdot, \cdot)$ denotes a binomial random variable, and $q_m(t; \tilde{m}, j)$ and $T$ are given in (8)-(9).

**Remark 5.** Because the readmission window is long and the readmission probability is not large (with an average of around 10-20%), each discharged patient’s contribution to the readmission arrival rate on any given day is small. This leads to a smoothing effect on the readmission arrival rate across discharge policies as long as the day-to-day discharge actions do not fluctuate too much; see Figure 15 in Appendix D.3 for some numerical evidence. To maintain the Markov property of the MDP and obtain insightful structural properties, we consider the readmission arrivals $\{A_k^{'m,0}\}$ as exogenous variables for the technical results. In the numerical study and implementations (Sections 7 and 8), we relax this assumption and let the readmission arrival depend on past discharge actions. To (approximately) solve the MDP there, we estimate the readmission arrival distributions from a static policy developed in Section 5.2 that gives analytical forms.

**One-period cost.** The current period cost function is composed of the ward occupancy congestion cost, $c_h(X_k)$, and the discharge cost, $c_d(D_k)$, which depends on the expected number of readmissions given the patients being discharged. A reasonable form for $c_h(X_k)$ and $c_d(D_k)$ is

$$c_h(X_k) = C \cdot \left( \sum_{m=1}^{M} \sum_{j=0}^{J} X_k^{m,j} - N \right)^+, \quad (12)$$

where $C$ is the unit holding cost and $\left( \sum_{m=1}^{M} \sum_{j=0}^{J} X_k^{m,j} - N \right)^+$ is the number of patients that cannot be accommodated in a ward bed at epoch $k$;

$$c_d(D_k) = \sum_{m=1}^{M} \sum_{j=0}^{J} \sum_{t=0}^{T} R_t \mathbb{E} \left[ \text{Bin}(D_k^{m,j}, q(t; m, j)) \right] = \sum_{m=1}^{M} \sum_{j=0}^{J} D_k^{m,j} \sum_{t=0}^{T} R_t q(t; m, j), \quad (13)$$
where \( q(t; m, j) \) is given in (8), and \( R_t \) denotes the corresponding penalty cost. We allow the penalty cost \( R_t \) to depend on the timing of readmission since early readmitters are found to require more intensive care than late readmitters (Skolarus et al. 2015). In the case where \( R_t = R \) for all \( t \), we have \( c_d(D_k) = \sum_{m=1}^{M} \sum_{j=0}^{J} D_k^{m,j} R \cdot r(m, j) \) from (8).

Note that cost parameters \( C \) and \( R_t \)’s may be difficult to estimate in practice. In our analysis we employ them primarily as “tuning parameters” to reflect the tradeoff between discharging too few versus too many patients. In the numerical study and the implementation, we derive efficient frontiers for the decision makers to identify an operating regime to achieve their desired performance measures, eliminating the reliance on the cost parameters themselves.

**Objective.** Let \( \Pi = \{ D_k^{m,j} : 0 \leq D_k^{m,j} \leq X_k^{m,j} \} \) be the set of admissible policies and let \( \{ X_k^{\pi} \} \) be the resulting state under policy \( \pi \). We can now write the objective function for the infinite-horizon long-run average cost problem for policy \( \pi \in \Pi \):

\[
Z^\pi = \limsup_{\tau \to \infty} \frac{1}{\tau} \sum_{t=0}^{\tau-1} \mathbb{E} \left[ c_h(X_k^{\pi}) + c_d(D_k) \right].
\] (14)

The optimal solution to (14) is therefore given by

\[
\gamma^* = \inf_{\pi \in \Pi} Z^\pi.
\] (15)

Theorem 1 proves the existence of an average cost optimal stationary policy, with proof detailed in Appendix B.1.

**Theorem 1.** For the average-cost optimality equation defined by Equations (14) and (15), there exists an average-cost optimal stationary policy.

**5. Bellman equation and structural properties**

In this section, we derive analytical properties of the long-run average cost MDP. Let \( A_k \) denote the vector of random arrivals from each class, including both new and readmission arrivals (see remark 5). For a given state \( x = (x^0, x^1, \ldots, x^J) \) with \( x^j = \{ x^{1,j}, \ldots, x^{M,j} \} \), the Bellman equation is given by

\[
V(x) = \min_{D \in \Pi} c_h(x) + c_d(D) - \gamma^* + \mathbb{E}_{A_k} V(A_k, x^0 - D^0, x^1 - D^1, \ldots, x^{J-1} - D^{J-1}) \quad \forall x \in \mathcal{S},
\] (16)

where \( D \) is a vector with component \( D^j = (D^{1,j}, \ldots, D^{M,j})' \), \( \mathcal{S} \) is the state space, \( \gamma^* \) is the optimal long-run average cost, and \( V(\cdot) \) is the (relative) value function. In Section 5.1, we analyze the structure of the optimal solution to the Bellman equation and establish two properties (1) a ranking of patients to discharge, and (2) a threshold discharge policy that follows the ranking. In Section 5.2 we analyze a special case of the MDP, where the action space is constrained to static-threshold policies. Leveraging the tractability of this special case, we investigate the interplay between system congestion and discharge risk. The analytical results developed in these two subsections provide insights into our two key research questions: “who to discharge” and “how many to discharge.”
5.1. Action space: structural properties for patient rank and discharge threshold

In this section, we show that patients can be ranked by their marginal risk by defining a property we call **strong dominance**. That is, the optimal discharge policy will discharge all patients of a higher rank before discharging any patients of a lower rank in terms of strong dominance. Strong dominance describes the shape of the patient risk trajectory as a function of LOS, \( r(m,j) \). At a high level, one trajectory strongly dominates another if the slope in the LOS dimension is smaller.

**Definition 1 (Strong Dominance).** Define patient type via (class, LOS). Type \((m_1,t_1)\) strongly dominates type \((m_2,t_2)\), or \((m_1,t_1) \succ (m_2,t_2)\), iff

\[
\frac{\partial r(m_1,t_1+t)}{\partial t} \leq \frac{\partial r(m_2,t_2+t)}{\partial t} \quad \forall t \geq 0.
\]  

Note, strong dominance always holds within a class \(m\). That is for any \(t_1 > t_0\) then \((m,t_0) \succ (m,t_1)\) because \(r(m,\cdot)\) is decreasing and convex; see proof of convexity in Appendix B.5. In the predicted risk trajectories using the dataset from our partner hospital, strong dominance is a complete order, i.e., it holds between all pairwise comparisons of patient type trajectories. Though our dataset suggests strong dominance always holds, for the sake of completeness and broader applicability of our analytical tool, in Appendix C we propose and analyze a more general **weak dominance** criterion.

The following proposition, with proof in Appendix B.2 demonstrates that keeping the strong dominant patient longer provides less benefit (smaller discharge cost reduction) than keeping the dominated patient longer. Let \(e_{(m,t)}\) be the unit vector with a 1 in the index corresponding to type \((m,t)\) and zero elsewhere. Adding this vector to \(D_k\) indicates adding a single discharge of patient type \((m,t)\) in the action.

**Proposition 1.** For \(c_d(D)\) of the form (13), and \(R_t = R\), if \((m_1,t_1) \succ (m_2,t_2)\), then in any epoch \(k\), the following holds for any future epoch, \(k' > k\):

\[
c_d(D_k + e_{(m_1,t_1)}) - c_d(D_{k'} + e_{(m_1,t_1+k'-k)}) \leq c_d(D_k + e_{(m_2,t_2)}) - c_d(D_{k'} + e_{(m_2,t_2+k'-k)})
\]

Next, we leverage Proposition 1 to prove a theorem that allows us to rank patients and discharge them in strict order of their ranking. We specify the proof for the simpler case where the two interchanged patients have not reached their maximum LOS, and leave the other more tedious case to Appendix B.3. Below, the unit vector \(e_{(m,t)}\) multiplying \(D_k\) or \(X_k\) gives the number of \((m,t)\) patients in the corresponding action or state, e.g., \(D_k \cdot e_{(m,t)} = D_k^{m,t}\).

**Theorem 2.** For \(c_d(D)\) of the form (13), and \(R_t = R\), consider two patient types, \((m_1,t_1) \succ (m_2,t_2)\). Then, the optimal discharge action \(D_k^* \cdot e_{(m_2,t_2)} > 0\) only if \(D_k^* \cdot e_{(m_1,t_1)} = X_k \cdot e_{(m_1,t_1)}\), i.e., we would discharge patient type \((m_2,t_2)\) only if we have discharged all type \((m_1,t_1)\) patients.

**Proof.** We prove the theorem via interchange argument on two patients of type \((m_1,t_1)\) and \((m_2,t_2)\), respectively. Suppose at period \(k\), that \(D_k^* \cdot e_{(m_2,t_2)} = 1\) and that \(D_k^* \cdot e_{(m_1,t_1)} = X_k \cdot e_{(m_1,t_1)} - 1\). Note, if \(D_k^* \cdot e_{(m_2,t_2)} = n\) and \(D_k^* \cdot e_{(m_1,t_1)} = X_k \cdot e_{(m_1,t_1)} - n\), we can repeat the interchange argument iteratively to achieve the same result because the discharge cost function is linear. Suppose this one
patient of class \((m_1, t_1)\) that was not discharged in period \(k\) is discharged at a later time, \(k'\). Call this policy \(\pi_1\), with value function \(V^\pi_1(X_k)\). Now consider a second policy, \(\pi_2\) that switches the discharge timing of the type \((m_1, t_1)\) patient and the type \((m_2, t_2)\) patient that we track in the interchange argument. All other actions remain the same. First, consider the case where \(t_1 + k' - k \leq J\), i.e. that \(k'\) is not beyond type \((m_2, t_2)\) patient’s maximum LOS. Let \(D_k\) and \(D_{k'}\) be the discharge actions excluding the two switched patients.

\[
V^{\pi_2} - V^{\pi_1} = c_d(D_k + e_{(m_1, t_1)}) + c_d(D_{k'} + e_{(m_2, t_2)}) - c_d(D_k + e_{(m_1, t_1 + k' - k)}) - \left( c_d(D_k + e_{(m_2, t_2)}) - c_d(D_{k'} + e_{(m_2, t_2 + k' - k)}) \right) \\
= c_d(D_k + e_{(m_1, t_1)}) - c_d(D_{k'}) - c_d(e_{(m_1, t_1 + k' - k)}) - \left( c_d(D_k + e_{(m_2, t_2)}) - c_d(D_{k'} + e_{(m_2, t_2 + k' - k)}) \right) \\
\leq 0.
\]

The first equality follows because the occupancies are the same under both policies under all sample paths, so the occupancy costs cancel out and so do the discharges costs for all patients except the two interchanged patients, since our discharge cost is linear. The inequality follows from Proposition 1 which shows that the first two terms in the second line are smaller than the second two. The proof for the second case where \(t_1 + k' - k > J\) is given in Appendix B.3. Since \(\pi_2\) produces a smaller cost than \(\pi_1\), \(\pi_1\) cannot be the optimal policy, contradicting the assumptions made at the beginning. □

Theorem 2 implicitly provides a ranking of patient types. In particular, this discharge ranking depends on marginal risk, not absolute risk as has been the prevailing approach according to our discussions with hospitals. We now formalize the ranking letting \([i]\) be the \(i^{th}\) ranked patient type, which means that \([j] > [i] \forall j < i\); \([1]\) being the highest rank (most “desirable” to discharge). We prove the univariate threshold structure of the optimal policy.

**Corollary 1.** The optimal discharge policy, \(D_k^*\) for the optimization defined by (16) is of threshold form, with univariate threshold \(\bar{D}^*(X)\), where

\[
D_k^* \cdot e_{[i]} = \min \left( X_k^{[i]}, \left( \bar{D}(X_k) - \sum_{j<i} X_k^{[j]} \right)^+ \right) \quad (18)
\]

**Proof.** We prove this by contradiction. Suppose the policy, \(D_k\), is not of the form in (18). Then it must be the case that there exists \(D_k \cdot e_{[i]} > 0\) and \(D_k \cdot e_{[j]} < X_k^{[j]}\) for some \(j < i\). However, this cannot be optimal by Theorem 2. □

**Remark 6.** For ease of exposition, in the rest of this paper we let \(R_t = R\). The analytical results from this section extend to the more tedious time-varying case \(R_t\) if we redefine the strong dominance criterion for \((m_1, t_1) \succ (m_2, t_2)\) as \(\frac{\partial \bar{R}(m_1, t_1 + l)}{\partial t} \leq \frac{\partial \bar{R}(m_2, t_2 + l)}{\partial t} \forall t \geq 0\) where \(\bar{R}(m, j) = \sum_{i=0}^{J} R_t q(m; j, t)\).

### 5.2. Special case: class-dependent, state-independent thresholds

In this section, we present a special case of the MDP problem motivated by our discussions with hospital managers, where the action space is restricted to class-dependent, but state-independent threshold policies, i.e., a patient is not discharged until her risk level drops below a pre-set threshold.
Because these thresholds do not change with respect to system state, we call the resulting policy a static policy. In addition to generating analytical insights, this static policy also provides a benchmark to compare against our dynamic policy developed in Section 6.

At each day \( k \), we discharge all class \( m \) patients whose risk is at or below threshold \( s_m \). Given the predicted risk trajectory for each class \( m \), there is a one to one relationship between \( s_m \) and the LOS. Let \( l_m \), be the LOS that corresponds a class \( m \) patient reaching risk threshold, \( s_m \). Optimizing the thresholds \( (s_1, \ldots, s_M) \) is equivalent to optimizing \( (l_1, \ldots, l_M) \).

For a given policy \( \pi \) with thresholds \( (l_1, \ldots, l_M) \), the distribution on number of discharges from class \( m \) on day \( k \), \( D^m_k = X^m_{k-l_m} \), is stationary under \( \pi \). Let \( l_m \) denote the steady-state expectation of \( D^m_k \), and \( E[Q] \) denote the expected queue length under policy \( \pi \) in the steady state. Letting \( R_t = R \), we show in Appendix B.6 that minimizing the long-run average cost is equivalent to minimizing

\[
R \cdot \sum_{m=1}^{M} l_m \cdot \sum_{j=1}^{m-1} q_m(t; \tilde{m}, l_{\tilde{m}}) \cdot E[D_{\tilde{m}}] + C \cdot E[Q].
\] (19)

\( E[D^m] \) can be found by solving the following set of flow-balance equations:

\[
E[D^m] = E[A^m_{k-l_m}] + E[A'_{k-l_m}] = \Lambda_m + \sum_{\tilde{m}=1}^{M} \sum_{t=1}^{T} q_m(t; \tilde{m}, l_{\tilde{m}}) \cdot E[D_{\tilde{m}}], \quad m = 1, \ldots, M.
\] (20)

If patients do not change classes upon readmission, we have \( E[D^m] = \Lambda_m/(1 - r(m, l_m)) \) by (8).

For \( E[Q] \), note we discharge all \( D^m_k = X^m_{k-l_m} \) patients on each day \( k \). \( X^m_{k-l_m} \) is composed of new arrivals, \( A^m_{k-l_m} \sim \text{Poiss}(\Lambda_m) \) and readmissions, \( A'_{k-l_m} \), which we approximate as a Poisson for tractability. Then,

\[
E[Q] \approx E \left[ \left( \sum_{m=1}^{M} \sum_{j=0}^{l_m-1} \text{Poiss} \left( \Lambda_m + E[A'_{k-l_m}] \right) - N \right)^+ \right].
\]

In Section 5.1 we found that the optimal policy discharges patients according to their ranking up to a certain threshold. In the following proposition, we characterize the threshold in this special MDP to provide insight into the question of “how many to discharge” via the optimal occupancy in steady-state. For ease of exposition, we present the results for a single class of patients; the insights extend to the multi-class problem. We drop the index on class \( m \) and use \( \ell \) to denote the discharge threshold, with a slight abuse of notation to define the associated discharge risk as \( r(\ell) \). Let \( \tilde{\Lambda} = \Lambda/(1 - r(\ell)) \) denote the total arrival rate including readmissions.

**Proposition 2.** Under a normal approximation for the Poisson distribution with mean \( \tilde{\Lambda} \), the optimal discharge threshold \( \ell \) solves the following equation

\[
R \cdot \left( \frac{d}{d\ell} r(\ell) + r'(\ell) \tilde{\Lambda} \right) = C(1 - \Phi(\alpha)) \frac{d B}{d\ell},
\] (21)

where \( \Phi(\cdot) \) is the CDF of standard normal, \( B = \ell \cdot \tilde{\Lambda} \) is the system offered load, and \( \alpha = \frac{N-B}{\sqrt{B}} \).
The proof is given in Appendix B.6. Equation (21) characterizes the optimal occupancy under the given cost parameters in terms of the marginal increase in readmission risk versus marginal increase in congestion. Note that the left-hand side depends on the marginal change in the discharge risk as a function of the threshold $\ell$. The right-hand side is the current probability of exceeding capacity, $(1 - \Phi(\alpha))$, times the marginal increase in system workload from keeping patients longer; i.e., the marginal for system congestion. Note that $1 - \Phi(\alpha) = P(Z > \alpha) = P\left(Z > \frac{N - B}{\sqrt{B}}\right) = P\left(B + Z\sqrt{B} > N\right)$, where $B + Z\sqrt{B}$ approximates the total number of patients (or workload) in system.

6. Dynamic discharge decision support

In this section, we develop an efficient dynamic algorithm to solve the optimization problem presented in Equations (14) and (15) under the full action space. Conventional MDP techniques such as value or policy iteration require solving the Bellman equation (16), which suffers from the curse of dimensionality. If we cap the number of patients of each type (class, LOS) to be $\bar{S}$, both the state space and action space are of size $\bar{S}^{M \cdot (J + 1)}$. Even in a simple two-class setting where patients are kept at most 3 days and $\bar{S} = 30$, the state space is of size $30^{4 \times 2} = 6.56 \times 10^{11}$. To overcome this challenge, we start from a special quadratic cost setting in Section 6.1, where we are able derive explicit solutions in a linear-quadratic form for the optimal action and value function. Then in Section 6.2, we leverage the linear-quadratic solution, along with the structure of the optimal policy obtained in Section 5.1, to transform the original MDP into a univariate optimization problem, which significantly reduces the computational complexity. In Section 6.3, we demonstrate the near-optimal performance of this dynamic algorithm in small scale problems where the conventional value iteration is still feasible.

6.1. Linear quadratic optimization

In this section, we consider a finite-horizon version of (14) with $T + 1$ periods and a quadratic structure for $c_h(\cdot)$ and $c_d(\cdot)$. Let $S_k = \sum_{m,j} X_{m,j}^k$ denote the total number of patients in the system, $c_h(X_k) = C \cdot S_k^2$, and $c_d(D_k) = R \cdot \sum_{m,j} r(m,j)(D_{m,j}^k)^2$. Furthermore, we assume that (i) $X_{m,j}^k$ is an absorbing state, i.e., once reaching day $J$ (the maximum LOS in the original problem), the patients will stay in that state, and (ii) the action space is unconstrained, i.e., $\Pi_{un} = \{D_{m,j}^k \in \mathbb{R}\}$. Given the initial state $X_0$, the optimization problem can be written as

$$V_{k,LQ}^\pi(X_k) = \sum_{i=k}^{T} \mathbb{E}\left[C \cdot S_i^2 + R \cdot \sum_{m,j} r(m,j)(D_{m,j}^k)^2\right] + \mathbb{E}\left[C \cdot S_{T+1}^2\right],$$

$$V_{LQ}^\pi(X_0) = \min_{\pi \in \Pi_{un}} V_{0,LQ}^\pi(X_0).$$

(22)

Proposition 3 shows that in this finite-horizon setting, the optimal discharge action is linear in $S_k$, and the value function is a quadratic function of $S_k$. 
**Proposition 3.** The optimal discharge decision of problem (22) is given by

\[ D_{k}^{m,j} = a_{k}^{m,j}S_{k} + b_{k}^{m,j}, \]  

and the value function is given by

\[ V_{k,LQ}^{\pi}(X_{k}) = \alpha_{k}S_{k}^{2} + \theta_{k}S_{k} + \kappa_{k}. \]  

Here, \( a_{k}^{m,j} \) denotes the \((j-1)M+m^{th}\) entry of vector \( U_{k}^{-1}(\alpha_{k+1}, \ldots, \alpha_{k+1})' \), and \( b_{k}^{m,j} \) denotes the \((j-1)M+m^{th}\) entry of vector \( U_{k}^{-1}B_{k} \). The matrix \( U_{k} \) and the vector \( B_{k} \) are given by

\[ U_{k} = \begin{pmatrix} \alpha_{k+1} + R_{0} & \alpha_{k+1} & \alpha_{k+1} & \ldots & \alpha_{k+1} \\ \alpha_{k+1} & \alpha_{k+1} + R_{1} & \ldots & \alpha_{k+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha_{k+1} & \alpha_{k+1} & \ldots & \alpha_{k+1} + R_{J} \end{pmatrix}, \quad B_{k} = \begin{pmatrix} \alpha_{k+1}\mathbb{E}[A_{k}] + \theta_{k+1}/2 \\ \alpha_{k+1}\mathbb{E}[A_{k}] + \theta_{k+1}/2 \\ \vdots \end{pmatrix}, \]

where \( R_{j} = R \cdot (r(1,j), \ldots, r(M,j))' \), \( A_{k} \) denotes the total arrivals from all classes including readmissions, and constants \( \alpha_{k}, \theta_{k} \) and \( \kappa_{k} \) can be recursively calculated using \( a_{k}, b_{k}, \alpha_{k+1}, \theta_{k+1}, \kappa_{k+1} \).

The proof in Appendix B.4 uses an induction argument where we also provide specifications of the constants \( \alpha_{k}, \theta_{k} \) and \( \kappa_{k} \). Equation (23) connects to the linear decision rule used in stochastic optimization (Chen et al. 2008). We use this linear decision rule in the next section as an approximation for the optimal action in future periods. The quadratic value function also provides an approximation for the cost-to-go. We refer to these approximations as the linear quadratic approximation.

### 6.2. Dynamic discharge algorithm

In this section, we develop an efficient dynamic heuristic that transforms the original MDP into a univariate optimization. This transformation relies on two key properties: (i) the structural properties proved in Section 5.1 map the high-dimensional action space into an equivalent univariate action space; and (ii) the approximations developed in Section 6.1 allow us to approximate the cost-to-go with a quadratic function of the total occupancy, which only depends on the univariate action in (i).

We first specify (i) and (ii), and then present the univariate optimization.

**Action space reduction.** We use Corollary 1 to reduce the action space to a univariate decision, \( \hat{D} = \hat{D}(X_{k}) \), the total number of discharges on day \( k \). From Equation (18) there is a one-to-one mapping from \( \hat{D} \) to the discharge action \( D_{k}^{m,j} = D_{k}^{m,j}(\hat{D}) \) for each class \( m \) and LOS \( j \).

**Cost-to-go approximation and univariate optimization.** To approximate the cost-to-go function (the last term in (16)), we start from the state in period \( k \), expand period \( k+1 \) with actions approximated by the linear decision rule (23), and use the quadratic approximation for the future cost (period \( k+2 \) and beyond) with \( V_{k,LQ}^{\pi}(S_{k+2}) \) given by (22). The optimal total discharges is therefore

\[ \hat{D}^{\pi}(X_{k}) = \arg\min_{0 \leq \hat{D} \leq S_{k}} \left\{ C(S_{k} - N)^{+} + \sum_{m,j} R \cdot r(m,j)D_{k}^{m,j}(\hat{D}) + \mathbb{E} \left[ C(S_{k+1} - N)^{+} + \sum_{m,j} R \cdot r(m,j)(a_{k+1}^{m,j}S_{k+1} + b_{k+1}^{m,j}) + \mathbb{E}[V_{k,LQ}^{\pi}(S_{k+2})] \right] \right\}. \]  

(25)
The first two terms capture the congestion cost and the discharge cost, for the current period, $k$. The third term captures the congestion cost in period $k+1$, while the fourth and fifth terms capture the discharge cost in period $k+1$. The fourth term is the discharge cost for patients who have reached their maximum LOS in period $k+1$, and the fifth term is the discharge cost for all other patients approximated by the optimal actions (23). The final term is the quadratic approximation for future cost in period $k+2$ and beyond.

To see why (25) is a univariate optimization, notice that $S_{k+1} = S_k - \tilde{D} + A_k$, and $S_{k+2} = S_{k+1} - \sum_m (X_{k}^{m,J-1} - D_{k}^{m,J-1}(\tilde{D})) - \sum_m \sum_{j \neq J} (2a_{k+1}^j S_{k+1} + b_{k+1}^j) + A_{k+1}$; both only depend on $\tilde{D}$ given the realization of the total arrivals (from all classes), $A_k$ and $A_{k+1}$, in periods $k$ and $k+1$. Further, $V_{LQ}^*(S_{k+2})$ only depends on the total occupancy, $S_{k+2}$. Thus, this value only depends on the total number of dischargers, i.e. the univariate decision variable $\tilde{D}$.

**Tuning parameters.** If the number of patients reaching the maximum LOS, $X_{k}^{m,J-1} - D_{k}^{m,J-1}(\tilde{D})$, is small, which we expect to be the case given $J$ is chosen as an upper bound on LOS, then $S_{k+2} \approx \left(1 - \sum_m \sum_{j \neq J} a_{k+1}^j S_{k+1} + \sum_m \sum_{j \neq J} b_{k+1}^j + A_{k+1}\right)$ is linear in $S_{k+1}$ and $A_{k+1}$. Recall from Proposition 3 that $V_{LQ}^*(S_{k+2}) = a_0 S_{k+2}^2 + \theta_0 S_{k+2} + \kappa_0$. Thus, $\mathbb{E}_{A_{k+1}}[V_{LQ}^*(S_{k+2})]$ can be written as a quadratic function in $S_{k+1}$. Using the linear structure of $(a_{k+1}^j S_{k+1} + b_{k+1}^j)$ in $S_{k+1}$ and taking out $C(S_k - N)$ since it does not depend on $D_k$, we can simplify (25) as

$$\hat{D}^*(X_k) = \arg \min_{0 \leq \hat{D} \leq S_k} \left\{ \sum_{m,j} R \cdot r(m,j) D_k^{m,j}(\hat{D}) + C(S_{k+1} - N)^+ + \mathbb{E}_{\tilde{A}_k} \left[ a_0 S_{k+1}^2 + \beta S_{k+1} + \kappa \right] \right\}, \quad (26)$$

where $\tilde{\alpha}$ and $\tilde{\beta}$ can be treated as tuning parameters. In the implementation of the dynamic algorithm, we fine tune $\tilde{\alpha}$ and $\tilde{\beta}$ to achieve a better performance (in cost minimization).

The quadratic approximation for cost-to-go based on $V_{LQ}^*(S_{k+2})$ works well in most of our settings because in our general formulation (12), the queue length $(S_k - N)^+$ is a piecewise linear function and can be approximated by quadratic functions in $S_k$.

**Remark 7 (Weak dominance.).** The one-to-one mapping (18) relies on strong dominance, which holds for risk trajectories predicted from our dataset. If strong dominance fails, we develop an alternate ranking called weak dominance, which is based on a Decomposition Heuristic 1 detailed in Appendix C.1. We find that the key factor driving the discharge quantity is a weighted average of the marginal improvements in readmission risk over the patient’s remaining LOS trajectory,

$$\omega(m,j) = \psi_1^m \cdot (R_{m,j} - R_{m,j+1}) + \psi_2^m \cdot (R_{m,j} - R_{m,j+2}) + \cdots + \psi_{j-j}^m(R_{m,j} - R_{m,j}) \sum_{t=1}^{J-j} \psi_t^m = 1, \quad (27)$$

where $R_{m,j} = R \cdot r(m,j)$. Note that strong dominance implies weak dominance.

### 6.3. Performance of the dynamic algorithm

In this section, we demonstrate the performance of the algorithm by (1) comparing the actions solved from the dynamic algorithm with those from value iteration in small-scale MDP’s and show that the dynamic algorithm is able to achieve near-optimal performance; and (2) comparing the dynamic
algorithm with several heuristics, including the static threshold policy and a one-step improvement on the static threshold policy, in a simulation setting for a suite of different parameters. In the interest of space, we demonstrate one set of results for the small-scale MDP below and leave other experiments to Appendix D: (i) additional numerical results in the small-scale MDP, including when the strong dominance no longer holds, and (ii) comparison with additional heuristics.

In the small-scale MDP, we consider two classes of patients with the maximum LOS $J = 3$ days. The risk trajectory for each class follows $r(1, j) = \{1, 0.3, 0.2, 0.15\}$ and $r(2, j) = \{1, 0.3, 0.08, 0.06\}$, satisfying the strong dominance criterion. We set $C = 1$, $R = 6$ and $N = 12, 20$. Figure 6 shows the gap between the optimal actions obtained from value iteration versus the actions obtained from our dynamic heuristic for each class. For comparison, we also plot the gap for a “myopic policy,” where we ignore the cost-to-go and set $\tilde{\alpha} = \tilde{\beta} = 0$ in (26). From all the states the maximum gap between our dynamic heuristic and the optimal action is one (see 6a) or two in only a few states (see 6b), whereas the myopic policy performs poorly. This demonstrates the importance of a more nuanced approach to discharge management.

7. Improving over practice: counterfactual and simulation analysis

In this section, we develop a case study based on data from our partner hospital. We begin with a counterfactual analysis on the hospital’s historical practice to demonstrate how our dynamic algorithm could have improved the hospitals performance; see Section 7.1. This has served as an important step in our implementation process to demonstrate to the hospital the potential value of our tool and to explain the logic for when our recommendations differed from their historical practice. In Section 7.2, we develop a data-driven discrete-event simulation to generate insights for the broader application of our methods in a wide range of hospitals through sensitivity analyses.

7.1. Trace driven counterfactual analysis based on historical data

To gain buy-in from management for a pilot implementation, we design a counterfactual to compare the dynamic policy (based on the dynamic algorithm) with the historical practice in our partner hospital. To create a realistic comparison with historical discharge behavior, we use a trace driven approach (Sherman and Browne 1973) in which the system inputs are generated from observations in
the data instead of using parametric assumptions. We then use an additive/subtractive counterfactual of readmission events to evaluate the impact of changing historical discharge decisions along several dimensions. Our results demonstrate (1) Pareto dominance of dynamic policy over the historical practice, (2) a high positive catch rate – properly identifying and intervening on patients who were readmitted in the data, and (3) occupancy smoothing – an unintended additional benefit.

**Avoiding/adding readmission events.** Our dynamic discharge policy may either extend or shorten the actual LOS observed in the data. If a patient had a readmission event in the data and our model recommends extending their LOS, we avoid the readmission with a probability that is proportional to the risk reduction from extending LOS. Specifically, we calculate the ratio between the predicted risk at our recommended discharge date versus the predicted risk at the actual discharge date. If a uniformly generated random variable exceeds this ratio, we avoid the readmission. To explain this approach, imagine that there is a random draw between 0 and 100% for each patient that produces the readmission event. If the readmission risk is 20% for this patient using the historical LOS, and we observe a readmission event in the data, then we know the realized outcome of the random draw lies anywhere between 0% and 20%. Suppose by extending the LOS, we reduce the readmission risk to 15%. Then we can avoid this readmission event if the realized random draw lies between 15% and 20%, with a probability \( \frac{5}{20} = 25\% \). Similarly, if a patient did not have a readmission event and our model recommends shortening her LOS, we generate a new readmission event with a probability proportional to the risk increase. We use Monte-Carlo simulation with 50 replications to generate the sequence of the random draws for each patient.

**Personalized risk curves.** In developing the modeling framework in Section 4, we group patients into classes to limit the state space for analytical tractability. However, in our implementation of the dynamic heuristic on real data, we relax this restriction to incorporate personalized risk trajectories from our prediction model, removing the reliance on patient classes. For the current period, we apply the strong dominance criterion (Definition 1) to rank all patients who are currently in the system. To evaluate cost-to-go for future periods, though, we still have to retain the classes and use the average risk curve for each class since we cannot know the individual characteristics of patients that have not yet arrived. In the case study, we cluster patients into \( M = 3 \) classes with the \( k \)-means method based on their predicted curves. Figure 13b in Appendix A.4 plots the average risk curve for each of the three classes, roughly corresponding to the low-, medium-, and high-risk groups, where the estimated proportion of each class is 48.6%, 33.4%, 18.0%, respectively.

**Tuning parameters.** Estimating the cost parameter \( C \) is challenging. Also in our data there is no reliable estimate of the capacity \( N \). Instead, we use these two as tuning parameters, and plot the efficient frontiers with respect to different performance metrics by varying \( N \) and the ratio between \( C \) and \( R \). This allows hospital managers to choose a parameter regime to achieve their desired performance metrics. During the actual implementation (detailed in Section 8), we use a “default setting” based on the management team’s feedback on their preferred target point of the efficient
Efficiency frontier. We set $N = 40$, $C = 1$ and change $R$ form 0.01 to 120 to get the solid line in each plot (using the dynamic heuristic). The red dot corresponds to the performance of the historical discharge policy. The 95% confidence intervals are tight and we omit them in the plots.

Summary of results. Figure 7 plots the readmission risk against the average LOS (which also implies the average occupancy by Little’s Law). In this figure, the solid blue line corresponds to the performance of dynamic policies from a series of experiments, where we vary $R$ from 0.01 to 120, with $N = 40$ and $C = 1$; the red dot corresponds to the performance of the hospital’s historical practice.

We report three types of readmission risk in each of the subplot:

- **Realized readmissions.** Historical number of readmissions (adjusted by the number of avoided or added readmissions for the dynamic policy) divided by the total number of patients.
- **Predicted readmissions.** To account for other sample paths than the realized one (in the data), we compare the predicted readmission risk using the historical vs the dynamic policies.
- **High and medium risk.** We compare the predicted readmission risk for the historical vs dynamic policies for only high and medium risk patients.

From the figure, the dynamic policy can either reduce the readmission risk while maintaining a similar average LOS, or maintain the same readmission risk with shorter LOS (lower occupancy); i.e. our dynamic heuristic exhibits Pareto dominance over the historical one. For example, the dynamic policy can significantly reduce readmission risk for medium- and high-risk groups (from 32% to 28%) when extending the LOS slightly (from 3.33 to 3.55 days). Table 2 reports the average added and avoided number of readmissions, along with the 95% confidence interval.

Positive catch. A *positive catch* is defined as extending the LOS for at least one day for a patient who was actually readmitted in the historical data. This performance metric measures the ability of our tool to properly identify and intervene on at-risk patients. It is also appealing to our industry partner (LCS) because they find that such metrics could be easily explained to hospital management. Using a policy that maintains a similar average LOS as the historical, we have a positive catch rate above 50%. Positive catch rates for other ratios of $R$ and $C$ are shown in Table 2.
### Table 2

Summary of statistics from dynamic policies under different cost parameters. Each number following the "±" sign in the first two rows denotes the half-width of the 95% confidence interval of the corresponding entry.

<table>
<thead>
<tr>
<th>R/C</th>
<th>0.01</th>
<th>0.1</th>
<th>1</th>
<th>3</th>
<th>40</th>
<th>120</th>
</tr>
</thead>
<tbody>
<tr>
<td># readm avoid</td>
<td>398±4.4</td>
<td>511±4.5</td>
<td>557±5.9</td>
<td>661±6.4</td>
<td>933±7.1</td>
<td>1475±7.4</td>
</tr>
<tr>
<td># readm add</td>
<td>518±7.1</td>
<td>382±6.2</td>
<td>338±6.1</td>
<td>263±4.6</td>
<td>124±2.9</td>
<td>25±1.6</td>
</tr>
<tr>
<td>Avg LOS</td>
<td>2.66</td>
<td>3.05</td>
<td>3.22</td>
<td>3.55</td>
<td>4.61</td>
<td>7.20</td>
</tr>
<tr>
<td>Positive catch</td>
<td>37%</td>
<td>45%</td>
<td>49%</td>
<td>54%</td>
<td>54%</td>
<td>68%</td>
</tr>
</tbody>
</table>

**Figure 8** Daily occupancy level on different days of a week. We set C = 1 and N = 40, and change R to get different average LOS under the dynamic policy (reported in the title of each subplot). The historical average LOS is 3.35 days.

**Occupancy smoothing.** Harrison et al. (2005) hypothesized that discharge policies could be effective in smoothing hospital occupancies, which has numerous benefits beyond the objectives of our study; such as reducing cancellations of elective surgery, boarding time in the emergency department, off-serving of patients, and stress on hospital staff, among others (Dai and Shi 2018, Helm and Van Oyen 2014, Kc and Terwiesch 2017). Even though occupancy smoothing is not explicitly incorporated in our objective function of the MDP, Figure 8 shows that the dynamic policy does produce a much smoother occupancy curve than the historical one – an unintended benefit. The peakedness of the occupancy (Bekker and Koeleman 2011), defined as the sum of squared differences between the daily and overall mean occupancy (normalized by the overall mean), reduces from 1.44 under the historical practice to less than 0.25 under the dynamic policy for settings in Figure 8.

### 7.2. Broader insights for different hospital environments and operational characteristics

In this section, we develop a high-fidelity, data-driven discrete event simulation to study the operational characteristics of hospitals that influence the efficacy of the dynamic discharge policy, where we vary the hospital unit size and utilization, shape of risk curves, and variability in the arrival process. Section 7.2.1 discusses the simulation design and parameterization. Section 7.2.2 identifies which types of hospitals benefit most from the dynamic policy through comparison with the static threshold policy and an empirical policy that mimics historical behavior. Section 7.2.3 highlights two operating regimes that we discovered: occupancy-driven versus quality-driven regimes.

#### 7.2.1. Simulation design and policy description

In this section, we introduce the simulation environment, two benchmark policies for comparison with the dynamic policy, and the parameter settings used in the simulation study.
High fidelity simulation. To create a high fidelity simulation, we incorporate a number of additional practical features on top of the modeling framework introduced in Section 4, including (i) using personalized individual patient risk curves; and (ii) adding a random delay for a discharged patient to leave their bed, which can be caused by paperwork after the roundings, pickup arrangement, etc. (Shi et al. 2015). In the baseline experiment, we assume the arrival rate is stationary; we vary the arrival rates by day of the week in Section 7.2.2.

Policy comparison. We compare the performance of the dynamic policy against two benchmark policies: (i) an empirical policy based on historical discharge behavior, and (ii) the optimized static threshold policy from Sec. 5.2. In the empirical policy, a patient is discharged when their risk reaches a class specific threshold with the three patient classes defined in the counterfactual in Section 7.1. Based on the average historical discharge behavior for each class, we estimate the discharge risk thresholds to be 10%, 20%, 35% for the low-, medium-, and high-risk class patients. The average LOS for each class is 3, 4, and 5 days with average LOS of 3.69 days, close to the historical average (3.35 days). We call this the empirical policy because it mimics the historical discharge behavior.

Baseline parameter setting and sensitivity analyses. For our baseline, we estimate parameters from our hospital dataset. The exogenous daily arrival rate is \( \Lambda = 6.5 \) patients. We select the cost parameters \( C = 1, R = 3 \) and capacity \( N = 40 \), under which setting the empirical policy has the smallest performance gap with the static and dynamic policies. In this way, we give the empirical policy the “benefit of doubt” by assuming that the hospital is aiming to optimize under the given system conditions. The personalized risk curve for each arrival is randomly drawn directly from predicted risk curves of all patients in the dataset. Patient readmission timing is generated similarly using the predicted Cox density curves.

To experiment with different load conditions, we vary \( N \) from 54 to 30, with the nominal utilization (under the empirical policy) ranging between 55% to 99.5%. To test the impact of hospital size, we increase the arrival rate to 9.75 and 13 (1.5 and 2 times of the baseline rate) and change \( N \) accordingly to achieve the same range of utilizations. To test the impact from the shape of the risk curves, we adjust the LOS coefficient in the prediction model to generate steeper/flatter curves. To study the impact of predictably variability in the arrival process, we test three day-of-week arrival patterns: a stationary arrival rate (no predictable variability), an empirical arrival pattern estimated from our hospital dataset, and a hypothetical arrival pattern with even larger variation.

7.2.2. Main insights for hospital operating characteristics

Figure 9 plots the performance improvement from the dynamic and static policies over the empirical policy (i.e., performance gap) under a variety of settings.

Size and utilization. Figures 9a to 9c show the performance under different utilizations when we increase the system size (reflected by the higher arrival rate while maintaining the same utilization). The performance of both dynamic and static policies exhibits a U-shaped pattern as a function of utilization, with the largest gains occurring at low and high occupancies. The more interesting finding
Figure 9  Performance gap under different sensitivity analysis settings. In the baseline, we set $\Lambda = 6.52$, $N = 40$, $C = 1$, and $R = 3$. To measure the utilization, we calculate the offered load using the empirical policy, since it does not react directly to occupancy levels. We then adjust the arrival rates to achieve different levels of utilization.

is that the gap between the static and dynamic policies converges as system utilization increases, and the gap converges faster as hospital size increases. We explain this phenomenon through the two operating regimes identified in Section 7.2.3. Thus, in smaller hospitals or single wards, and in hospitals with lower utilizations (common in community hospitals), the dynamic policy is needed to obtain the greatest benefit. In larger and/or more highly utilized hospitals, like urban and teaching hospitals, the simpler static policy may be sufficient. This insight may explain why the dynamic policy works well in our partner hospital, which is a small community hospital.

Risk curves. Figures 9d to 9f show the performance improvement for different shapes of the risk curves, moving from flatter to steeper slopes. As the slope increases, the gap between the dynamic and static policies increases significantly. The performance of the dynamic policy always improves as the slope increases, whereas the static policy does not exhibit this monotone behavior. Linking
Figure 10  Box plots for the post-discharge occupancy and average LOS. Post-discharge occupancy equals the current occupancy minus the number of discharges, divided by the capacity $N$. We set $\Lambda = 6.52$, $C = 1$, $R = 3$, and $N = 54$ or 25.

this to practice, the dynamic policy is more useful when patients recover faster. An example of this environment could be specialized elective surgery hospitals or wards, since these patients often have a more rapid recovery process than patients with complicated medical conditions.

**Day-of-week variability.** Figures 9g to 9i show the performance improvement when the arrival process exhibits an increasing day-of-week arrival variability. As the variability in the arrival process increases, the gap between dynamic and static increases significantly. The dynamic policy always performs better as variability increases, while the static policy does not necessarily improve.

The other interesting finding is that the dynamic policy helps smooth the daily occupancy similarly to the results of Section 7.1, whereas the static policy, even though the thresholds are optimized, exhibits little to no smoothing behavior; see Figure 16 in Appendix D. Hence, the dynamic policy is increasingly valuable along multiple dimensions as the variability in the arrival process increases, providing both better performance and greater occupancy smoothing. This again indicates that the dynamic policy may be more useful in elective surgical hospitals or wards, since surgical scheduling generates much of the variability in the arrival process.

7.2.3. Two operating regimes: congestion vs quality

In this section, we identify the key factors that drive the performance of the dynamic discharge policy demonstrated through the above sensitivity analyses. By analyzing the discharge actions of the dynamic policy under these different scenarios, we can classify the behavior as either occupancy-driven or quality-driven, where we appropriate the term quality to mean quality of care, which is different than the traditional definition in heavy-traffic queueing analysis (Gans et al. 2003).

**Occupancy-driven regime.** In this regime, typically in hospitals with low utilization, the discharge actions are more responsive to occupancy, approximately following an occupancy-based “discharge-up-to” policy, i.e., the post-discharge occupancy $X_k - D_k$ each day is roughly the same. Figure 10a shows a box plot of each day’s occupancy in the simulation. The whiskers are the extreme points and the blue lines represent the 25th and 75th percentile. When the utilization is low ($N = 54$, corresponding to 55% utilization under the empirical policy), the post-discharge occupancy exhibits little variation from day to day. The span of the extreme points (distance between the whiskers) is
smaller than 10%. The dynamic policy achieves such stable occupancy by keeping patients longer when occupancy is lower and shorter when occupancy is higher. This leads to greater variation in LOS, as illustrated the box plot for LOS in in Figure 10b; see $N = 54$ (lo) and $N = 54$ (hi) for plots of the low- and high-risk patients, respectively.

To explain this regime, notice that when the system is lightly loaded, there is more slack capacity to extend patients’ LOS to achieve a “preferred” discharge risk level. Essentially, the patients are allowed to stay until there is little change in discharge risk by extending or shortening LOS; i.e. where the risk curve is relatively flat. For example, the discharge risk only changes from 29% to 26% by extending from 9 and 12 days for high-risk patients; on average less than 1% change in risk per additional day. When the discharge risk becomes insensitive to LOS, the main driver of the discharge decision becomes the occupancy level. This is why the dynamic policy gains a more dominant performance improvement over the static policy since the latter does not respond to the day-to-day variations in occupancy. This phenomenon is accentuated by day-of-week variability, where the larger variation in the occupancy level rewards the greater flexibility afforded by the dynamic policy.

**Quality-driven regime.** In this regime, typically in hospitals with high utilization, discharge actions are less responsive to occupancy, instead following an approximate risk-based discharge-up-to policy; i.e., discharging patients up to a certain risk level. Unlike the high LOS variation in Figure 10b ($N = 54$), $N = 25$ (lo) and (hi) show that the quality-driven regime exhibits little LOS variation, demonstrating the risk-based threshold behavior. For example, for high-risk patients, the median LOS is 4 days, while the $25^{th}$ and $75^{th}$ percentiles are 3 and 4 days, respectively. Meanwhile, the occupancy has much larger variation; see Figure 10a, $N = 25$.

When the system is heavily loaded, there is relatively little flexibility to extend LOS since the policy seeks to discharge patients as soon as possible. Because patient LOS is generally short, the readmission risk is very sensitive to changes in LOS and the penalty for further shortening LOS outweighs the benefits gained by occupancy reduction due to the steep slope of the risk curve in this region. Thus, the “optimal” LOS remains similar regardless of the occupancy, leading to a policy that looks very similar to the static threshold policy. This explains the convergence of the preformance of the static and dynamic discharge policies at high utilizations.

8. Tool implementation in our partner hospital

*Through constant interaction with multiple doctors, case management and hospital executives, especially Chief Medical Officers, from community hospitals to academic hospitals, we have identified the need for discharge optimization as a valuable addition to the Readmissions Reduction offering of our tool.* – CEO of LCS.

In this section, we describe the process of working with LCS and our partner hospital to pilot the risk prediction and discharge optimization tool. The implementation process followed three phases. In phase one, our team met with the hospital management team, including the CEO and CMO. This meeting was facilitated by the fact that clinicians had asked LCS to develop new analytics to
provide a tool capable of leveraging inpatient intervention to reduce readmissions. After the meeting we received a commitment from management to support the pilot implementation process from both the IT side (phase two) and the personnel workflow side (phase three). We detail this process to provide guidelines for hospitals considering implementations of similar analytics tools.

8.1. Phase one: garnering management support for pilot implementation

Phase one of implementation involved validating the tool and providing proof of value. The report we delivered to hospital management included much of the analysis from Section 7 with a particular focus on the counterfactual, demonstrating how our method could improve over current practice. We also demonstrated the functionality of a pilot version of the web-based tool shown in Figures 3. Top selling points for the hospital were the positive catch rate (indicating accuracy of identifying targeted interventions) and the capability of differentiating early and late readmitters through our new clustering method. At this point, management requested a tiered implementation, starting with the prediction model to enable the clinicians to become comfortable with the new workflow, then adding the discharge recommendations on top of the existing platform.

8.2. Phase two: working with hospital IT and tool launch

Our prediction and optimization algorithms form the backbone of a cloud-based tool detailed below. To integrate this tool into daily operations, we develop the following operating procedure (summarized in Figure 11) for populating our models with data from the hospital’s IT system.

**Pilot Preparation.** To prepare for the pilot launch, we provided a template excel file to the hospital IT department describing an example of data features. This file shows a desired format for the data dumps with a list of patient characteristics needed for our models. Iterations were required since the hospital has additional data beyond what was requested. This process took about four weeks.

![Figure 11](image_url) Daily operations of the integrated tool for hospital use.

**Tool Launch.** LCS maintains a historical data archive, which can be used to re-train/re-parameterize the tool when necessary. During the daily operation of the tool, the team receives a rolling data dump each day, which contains all the past admissions/discharge and patient characteristic data from 60 days prior up to the present. The hospital transfers the daily dumps to the
LCS server one or two hours after midnight. The data is then combined with the historical archive to provide a full data set for our tool (after eliminating double entries during the merge). This data transfer is automatic, requiring no extra effort from the hospital.

**User interface.** Figure 3(a) in the introduction shows the discharge tool’s user interface. On the left side of the “Pre-discharge” module, every inpatient is represented by a green rectangle, which shows admission date and current LOS. Clicking on one of the patient rectangles reveals additional information such as principal diagnostic and relevant psycho-social data in the center panel. By clicking on “Submit” in the lower right hand corner, a window pops up with the discharge risk versus LOS curve shown in Figure 3(b). The vertical bars show the current LOS and the recommended discharge date from our optimization tool. Enlarged figures are available in Appendix F. At the hospital’s request, the tool also provides a post-discharge tab on the prediction of the time to readmission when discharge is initiated; see a snapshot in Figure 12(a) and enlarged plot in Appendix F.

In the pre-discharge module, the discharge recommendations are provided based on a default setting for the parameters $N$, $C$ and $R$, which were chosen by first determining which combination matches the historical average occupancy and LOS, and then soliciting management team’s input on their target occupancy level if it differs from the historical. We also provide a pop-up box in the tool, where a gatekeeper can change the default setting to allow for dynamic adjustment of management goals in the event, for example, that management has additional information or strategy that would require a temporary or more permanent change in policy. In this pop-up box, we demonstrate a “tradeoff” curve, where we vary the tuning parameters and demonstrate the tradeoff between average LOS (occupancy) versus average readmission rate; see Figure 12(b) and enlarged plot in Appendix F.

**Integration into current practice.** The tool is intended for decision support and can adapt to deviation from recommendations based on clinicians’ assessment and medical judgment. These deviations are logged in the tool, where clinicians can also provide notes about their discharge decision. The dynamic data gathered provides live feedback to our algorithms to make the necessary adjustments for more accurate/live predictions and decision support. Finally, the tool monitors performance over...
time, and notifies the LCS team if performance drops out of a predefined control region. When this occurs, the team initiates a re-parameterization of the model with more recent data. Historically, this has been done roughly every 6 months.

**Evidence of improvement.** For the pilot, we implemented the tool in the largest medical ward in our partner hospital as a test case. Below, we provide supporting evidence of the benefits from implementing stage 1, the discharge risk prediction, since this stage was implemented first (launched on November 7, 2017) and has been in operation long enough to gather data on hospital metrics. The discharge optimization part was put in prototyping (before official launch) on May 31, 2018, with testing and modifications still in progress when this paper is submitted.

This risk prediction tool is used both for discharge and post-discharge decision support. For the post-discharge, the hospital uses an internal metric called true-positive rate (TPR) as one way to monitor the performance of the risk prediction. TPR essentially measures the proportion of readmissions that could have been potentially prevented if a followup call were made 2-3 days before the actual readmission. This internal metric has improved from an average of 44.49% before implementation (from January 1 to September 30, 2017) to an average of 51.25% after implementation (from November 8, 2017 to March 31, 2018), which we believe has benefited from our new prediction tool.

So far the pilot has been well-received and is considered a success by our industry partners. Moving forward, LCS’ CEO indicates that: “Our tool has received very positive interest from another leading academic hospital on the east coast, praising the applicability of the real world decisions that need to be made on the ground and the simplicity of information that the tool communicates to the users. We are looking to roll out this module after more testing as part of the Readmissions Reduction Tool offering to all our current and future clients.”

8.3. **Phase three: hospital personnel training and patient education**

Integration into workflow is a two-way interaction that starts with soliciting user feedback to improve the tool interface and functionality in the pre-pilot phase. Prior to the official launch, LCS provides users a 30-60 minute training session, which is usually sufficient due to the simplicity of the user interface. The tool itself has a feedback mechanism where users can send requests to the LCS server for modifications, improvements, and fixes based on their experience with the tool.

The information provided by the tool on risk factors and data visualization of the evolving dynamics of a patient’s condition is intended to enable providers to communicate more effectively with their patients. For example, from the readmission risk curves provided by the tool, providers can educate the patient about when they may be at greatest risk for a readmission and effectively schedule a follow-up visit. Raising patient awareness may increase patient likelihood to self-report potential readmission causing conditions and improve compliance based on a better understanding of the risk. Using the risk factor data in the prediction model, the provider may also be able to educate the patient about their highest personal risk factors for readmission (e.g. anxiety, drug compliance, smoking or alcohol habits) in the hopes of increasing compliance.
9. Conclusion

In this paper, we develop a practical tool that integrates personalized readmission risk prediction into inpatient discharge planning. We develop, test, and implement this tool through collaborations with a data analytics company and a local partner hospital. Based on extensive counterfactual and simulation analyses, we demonstrate the value of this tool compared to the hospital’s historical discharge behavior and identify hospital characteristics that would benefit the most from our discharge optimization. We show that by increasing the average LOS moderately, the readmission risk can be reduced significantly, e.g., when the average LOS increases from 3.33 under the historical practice to 3.55 days under the dynamic policy, the corresponding readmission risk for medium- and high-risk groups reduces from 32% to 28%. Note that the purpose of our tool is not to promote early discharge (unless necessary when system is highly congested); rather it provides analytical support for hospital to balance the benefits of shortening or extending LOS.

This paper can be extended in a few directions. As mentioned in Section 4.2, the current dataset lacks entries to parameterize the extended model where patients experience different stages in the recovery process. We are working with the hospitals to explore the possibility of collecting this type of data. This would enable us to include time-varying covariates in the prediction model, which could help improve both risk prediction and discharge decision optimization. We focus on readmission risk as the main outcome metric, while future work could extend our framework to other patient outcomes. Newly emerging machine learning tools such as deep neural networks could also further improve the accuracy of the risk prediction tool. From the analytical side, we propose a weak dominance ranking criterion when strong dominance is not met, but not as a main focus of this paper. A more thorough study could take a deeper look into this setting along with establishing possible performance bounds, which likely would require new methodology and approximation methods.

References


Appendix A: Additional details for the readmission prediction model

A.1. Overview of the basic Cox model

Cox model (Cox 1992) is commonly used in modeling time-to-failure event, i.e., readmission in our setting. For a patient \(i\) with a baseline hazard rate function, \(h_0(\cdot)\), and a regression component, \(Y_i\beta\), her proportional hazard rate as a function of time, \(t\), is given by:

\[
h(t; i) = h_0(t) \exp(Y_i\beta),
\]

where \(h_0(t)\) captures the time component of the readmission event, and \(Y_i\beta = \sum_{k=1}^K \beta_k Y_{i,k}\) captures the dependence on patient-specific risk factors, \(Y_i\). A common choice for the baseline hazard function is the Weibull parametric form, i.e.,

\[
h_0(t) = \frac{k}{\lambda} \left(\frac{t}{\lambda}\right)^{k-1}, \quad \int_0^t h_0(u)du = \left(\frac{t}{\lambda}\right)^k.
\]

The parameters \(\lambda, k\) as well as \(\beta = \{\beta_1, \ldots, \beta_K\}\) need to be estimated.

With the hazard rate function \(h(t; i)\), the probability density of a readmission event occurring at time \(t\) for patient \(i\) is given by

\[
f(t; i) = h(t; i) \cdot \exp\left(-\int_0^t h(u; i)du\right).
\]

and the cumulative distribution function is

\[
F(t; i) = 1 - \exp\left(-\int_0^t h(u; i)du\right).
\]

The readmission time is usually available in the date, but not to a detailed hour. If a patient \(i\) is readmitted on day \(\tau_i\), the actual readmission time could lie anywhere between time \(\tau_i\) and time \(\tau_i + 1\). Thus, the corresponding probability is captured by

\[
F(\tau_i + 1; i) - F(\tau_i; i) = \left(1 - \exp\left(-\int_{\tau_i}^{\tau_i+1} h(u; i)du\right)\right) \cdot \exp\left(-\int_0^{\tau_i} h(u; i)du\right).
\]

Here, the first term \(1 - \exp\left(-\int_{\tau_i}^{\tau_i+1} h(u)du\right)\) can be interpreted as a discrete hazard for interval \([\tau_i, \tau_i + 1]\), i.e., the conditional probability that individual \(i\) will be readmitted in this interval given that she did not fail at the start of the interval. In other words, it is the discrete counterpart of \(h(t)\) in the continuous-time setting. Thus, \(F(\tau_i + 1) - F(\tau_i)\) can also be seen as the discrete version of \(f(t; i)\), the density function in the continuous-time setting.

A.2. EM algorithm

As detailed in Section 3.1, both the cured status \(C_i \in \{0,1\}\) and the membership variable \(Z_i \in \{1, \ldots, W\}\) for patient \(i\) are latent variables; i.e. unobservable. To perform parameter estimation with these latent variables, we use an expectation maximization (EM) framework.
A.2.1. Likelihood function with latent variables  Before we describe the EM algorithm, we first specify the individual and population-level likelihood function. To start, note we observe two types of patient in the dataset:

- The probability of observing a patient with no readmission event:
  \[ P(C_i = 1) + \sum_w P(C_i = 0, Z_i = w) \exp \left( - \int_0^{T_i} h^w(u; i) du \right), \]
  \[ (32) \]
  where \( T_i \) is the censored time for patient \( i \).

- If patient \( i \) is readmitted at time \( t_i \), then she is uncured. The probability of observing such a patient is
  \[ \sum_w P(C_i = 0, Z_i = w) h^w(t; i) \exp \left( - \int_0^{t_i} h^w(u; i) du \right). \]
  \[ (33) \]

As discussed in Appendix A.1, the readmission times are typically available in discrete time intervals (days). Hence, if a patient \( i \)'s readmission day is \( \tau_i \), then the probability of her readmission between \( \tau_i \) and \( \tau_i + 1 \) is given by:

\[ \sum_w P(C_i = 0, Z_i = w) \left( H^w(\tau_i + 1; i) - H^w(\tau_i; i) \right), \]
\[ (34) \]

where \( H^w(t; i) = 1 - \exp \left( - \int_0^t h^w(u; i) du \right) \) is the cumulative distribution function for cluster \( w \).

Now we are ready to specify the complete-data individual-level likelihood function for patient \( i \). Let \( J_i \) denote the number of readmission events for patient \( i \). For a patient \( i \) without a readmission event in the data (\( J_i = 0 \)), let \( T_i \) denote the censoring time (i.e., before \( T_i \) no readmission is observed), combining (32), (1), and (2), we have that

\[ L_i(T_i) = (\theta_{w_i})^{1\{C_i=1\}} \cdot (1 - \theta_{w_i})^{1\{C_i=0\}} \]
\[ \cdot \prod_{w=1}^W \left\{ \pi_w \cdot \exp \left( - \int_0^{T_i} h^w(u; i) du \right) \right\}^{1\{C_i=0, Z_i=w\}}, \quad J_i = 0. \]

To explain, the first line of \( L_i(t_i) \) captures the likelihood of being cured or not (stage 1), while the second line captures the likelihood of belonging each uncured group \( m \) and the readmission time being censored (stage 2).

Similarly, for a patient \( i \) who has been readmitted (\( J_i \neq 0 \)), let \( \tau_i \) be the day of readmission. Combining (34), (1), and (2), the complete-data individual-level likelihood function for patient \( i \) is given by

\[ L_i(\tau_i) = (\theta_{w_i})^{1\{C_i=1\}} \cdot (1 - \theta_{w_i})^{1\{C_i=0\}} \]
\[ \cdot \prod_{w=1}^W \left\{ \pi_w \cdot (H^w(\tau_i + 1; i) - H^w(\tau_i; i)) \right\}^{1\{C_i=0, Z_i=w\}}, \quad J_i \neq 0. \]
Using the two individual likelihood functions \( L_i(T_i) \) and \( L_i(\tau_i) \), we get the population-level likelihood function as given in Equation (4) in the main paper. The corresponding log likelihood function follows

\[
LL = \sum_{i} \left( \mathbb{1}\{C_i = 1\} \log \theta_{0i} + \mathbb{1}\{C_i = 0\}(1 - \log \theta_{0i}) \right)
+ \sum_{i \in J_{>0}} \sum_{w=1}^{W} \mathbb{1}\{C_i = 0, Z_i = w\} \left( \log \pi_w - \int_{0}^{T_i} h^w(u; i)du \right)
+ \sum_{i \in J_{=0}} \sum_{w=1}^{W} \mathbb{1}\{C_i = 0, Z_i = w\} \left( \log \pi_w + \log \left( H^w(\tau_i + 1; i) - H^w(\tau_i; i) \right) \right),
\]

(35)

**A.2.2. An Expectation Maximization (EM) Algorithm**

In each iteration of the EM algorithm, we first perform an E-step, and then perform an M-step.

**E-step:**

At the \( n^{th} \) iteration of the algorithm, we update the probabilities of \( P_{0,i}^{(n)} \) and \( T_{w,i}^{(n)} \) by calculating the posterior distributions of \( C_i \) and \( \{Z_i\} \) using the current estimation of the parameters, \( \Psi^{(n-1)} \).

* Cure probability update. For each patient \( i \) with \( J_i > 0 \), we know for sure the patient is uncured, and thus, we always have \( P_{0,i}^{(n)} = 0 \) in each iteration \( n \). For each patient \( i \) with \( J_i = 0 \), we perform a Bayesian update on the cure probability as

\[
P_{0,i}^{(n)} = \frac{\theta_{0i}^{(n-1)}}{\theta_{0i}^{(n-1)} + (1 - \theta_{0i}^{(n-1)}) \left( \sum_{w=1}^{W} \pi_w^{(n-1)} \cdot \exp \left( - \int_{0}^{T_i} h^w(u; i)du \right) \right)}.
\]

* Membership probability update. For each patient \( i \) with \( J_i > 0 \), her membership probability \( T_{w,i}^{(n)} \) is updated as

\[
T_{w,i}^{(n)} = \frac{\pi_w^{(n-1)} \left( H^w(\tau_i + 1; i) - H^w(\tau_i; i) \right)}{\sum_{w=1}^{W} \pi_w^{(n-1)} \left( H^w(\tau_i + 1; i) - H^w(\tau_i; i) \right)}, \quad w = 1, \ldots, W.
\]

(36)

For each patient \( i \) with \( J_i = 0 \), her membership probability \( T_{w,i}^{(n)} \) is updated as

\[
T_{w,i}^{(n)} = \frac{\pi_w^{(n-1)} \cdot \exp \left( - \int_{0}^{T_i} h^w(u; i)du \right)}{\sum_{w=1}^{W} \pi_w^{(n-1)} \cdot \exp \left( - \int_{0}^{T_i} h^w(u; i)du \right)}, \quad w = 1, \ldots, W.
\]

(37)

Finally, given the updated \( P_{0,i}^{(n)} \) and \( T_{w,i}^{(n)} \), we approximate the expected log likelihood function in the \( n^{th} \) iteration as

\[
\mathbb{E}[LL] \approx \sum_{i} \left( P_{0,i}^{(n)} \log \theta_{0i} + (1 - P_{0,i}^{(n)}) \log(1 - \theta_{0i}) \right)
+ \sum_{i} (1 - P_{0,i}^{(n)}) \sum_{w=1}^{W} T_{w,i}^{(n)} \log \pi_w
+ \sum_{i} (1 - P_{0,i}^{(n)}) \sum_{w=1}^{W} T_{w,i}^{(n)} \left( - \mathbb{1}\{J_i = 0\} \int_{0}^{T_i} h^w(u; i)du + \mathbb{1}\{J_i > 0\} \log \left( H^w(\tau_i + 1; i) - H^w(\tau_i; i) \right) \right).
\]

(38-40)
M-step:
In the $n^{th}$ iteration of the M-step, we obtain new estimation of $\Psi^{(n)}$ by maximizing the above expected log likelihood function. A nice feature of our approximate expected log likelihood function is that the coefficient $\xi$ in the cure stage (Stage 1, Figure 1), the mixing weights $\{\pi_w\}$, and the parameters for the Cox model in Stage 2 ($\{\lambda^w\}, \{k^w\}, \{\beta^w\}$) are decoupled, i.e., they appear separately in (38) to (40), respectively, and we can maximize each of these three parts independently.

To solve these maximization problems more efficiently, we adopt the so-called hard assignment, which converts the “soft labels” $P_{0,i}^{(n)}$ and $T_{w,i}^{(n)}$ into “hard labels” for each patient $i$. That is, in the $n$th iteration, patient $i$ is labeled as cured, or say, $\hat{C}_i^{(n)} = 1$ if $P_{0,i}^{(n)} \geq 0.5$, and uncured otherwise. If labeled as uncured, the patient is further assigned to cluster $\hat{w}_i$ by choosing $\hat{w}_i = \arg\max_{w=1}^{W} T_{w,i}^{(n)}$.

Then, maximizing (38) to get $\xi$ is equivalent to a logistic regression with labels $\{\hat{C}_i\}$, note that, because of decoupling and hard assignment, maximizing (40) is equivalent to solving the following problem for each $w$, using the subset of patients with label $\hat{w}_i = w$:

$$\max_{\lambda^w,k^w,\beta^w} \sum_{i \in \{\hat{w}_i = w\}} \left( \mathbb{1}_{\{J_i > 0\}} \log \left( H^w(\tau_{i+1};i) - H^w(\tau_i;i) \right) - \mathbb{1}_{\{J_i = 0\}} \int_{0}^{T_i} h^w(u;i) du \right).$$

(41)

To further specify the first term within the parenthesis of (41), let $\phi_i^w = \exp(Y_i^w_0\beta^w)$ for a given cluster $w$. From (31), we get

$$\log \left( H^w(\tau_{i+1};i) - H^w(\tau_i;i) \right) = \log \left( 1 - \exp \left( - \int_{\tau_i}^{\tau_{i+1}} h^w(u) du \right) \right) - \int_{0}^{T_i} h^w(u) du$$

$$= \log \left( 1 - \exp \left( - \phi_i^w \int_{\tau_i}^{\tau_{i+1}} h^w_0(u) du \right) \right) - \phi_i^w \int_{0}^{T_i} h^w_0(u) du.$$

Using the approximation for log function $\log(1 - x) \approx -x$, or equivalently, $1 - x \approx \exp(-x)$, we have

$$\log \left( 1 - \exp \left( - \phi_i^w \int_{\tau_i}^{\tau_{i+1}} h^w_0(u) du \right) \right) \approx \log \left( 1 - (1 - \phi_i^w) \int_{\tau_i}^{\tau_{i+1}} h^w_0(u) du \right)$$

$$= \log \left( \phi_i^w \int_{\tau_i}^{\tau_{i+1}} h^w_0(u) du \right).$$

To get the maximum likelihood estimate for $\beta^w$, $\lambda^w$, and $k^w$ from (41), we work with the negative log-likelihood function and use the optimization package in MATLAB on finding minimum of unconstrained multivariable function, which is based on the well-known BFGS Quasi-Newton method.

A.3. Details on fine-tuning the prediction model

Bootstrapping approach. To avoid overfitting, we use the bootstrapping approach as introduced in the main paper. In the conventional bootstrapping method, one randomly samples $n$ patients from the original data set with replacement where $n$ is the size of the original data. Then the prediction model is estimated with the boostrapped sample and the performance is tested on the original data set. In this paper, we use a sub-sampling bootstrapping method following the suggestion of De Bin.
et al. (2016) where we sample $m \approx 0.623n$ patients from the original data set without replacement for the training set. Then, we test the performance on the original data set.

**Number of clusters and initialization of EM.** Both the cured status and group assignment need to be initialized when performing the EM algorithm. To initialize $C_i$ for each patient at iteration 0, we label patients who are readmitted within 90 days as uncured ($C_i^{(0)} = 0$), and cured ($C_i^{(0)} = 1$) if they are not readmitted within 90 days. To initialize the group assignment $Z_i$ for patients who are labeled as uncured, we use the readmission timing as a partition criterion, i.e.,

$$Z_i^{(0)} = \begin{cases} 1, & \tau_i \in (0, t_1]; \\ 2, & \tau_i \in (t_1, t_2]; \\ \vdots \\ W, & \tau_i \in (t_{M-1}, t_M], \end{cases}$$

where $\tau_i$ is the $i$th patient’s readmission time, $t_j$ are post-discharge timeline partitions, and $W$ is the number of clusters. We test $W = 2, 3, 4$ and set $t_W = 90$ days to be consistent with the initialization cured cutoff. For $W = 2$, we test the following choices of $t_1$: 10, 15, 20, 25, 30, and 45 days. For $W = 3$, we test the following choices of $(t_1, t_2)$: (10, 30), (15, 30), (20, 30), and (30, 45). For $W = 4$, we test the following choices of $(t_1, t_2, t_3)$: (10, 30, 45), (15, 30, 45), (10, 30, 60), and (15, 30, 60). In the final implementation, we choose $W = 2$ with $t_1 = 30$ given its overall good AUC performance and robustness in parameter estimation. Note that more clusters do not necessarily mean a better performance because the number of readmissions are limited, and more clusters means fewer samples in one cluster.

**Selection of IV.** As mentioned in Section 3.3 of the main paper, we consider the following possible choices of IVs to address the endogeneity issue in LOS: admission day-of-week indicators, discharge day-of-week indicators, average occupancy levels on the day of discharge. Through different combinations, we find that the binary indicator of weekend versus weekday admission gives us the best average AUC performance.

**A.4. Prediction results**

For implementation, we choose our final model based on a combination of the best average AUC performance and the AUC obtained from the averaged estimate of coefficients $\overline{\Psi}$ from 50 bootstrap samples. The general recommendations for the bootstrapping approach is to use the largest number of samples that can be solved with reasonable level of computational effort. In our experiments, we found 50 samples to be a good balance between sample size and computation time. Figure 13a shows the predicted 90-day readmission cumulative probability, as a function of LOS, for 50 patients randomly selected from our dataset. Figure 13b shows the average readmission curves for 3 groups, which come from using $k$-means method to group patients based on their individual risk curves.
Appendix B: Major proofs

B.1. Proof of Theorem 1

Theorem 1. For the average-cost optimality equation defined by Equations (14) and (15), there exists an average-cost optimal stationary policy.

Proof. SEN assumptions from Sennott (2009) guarantee that the theorem holds. To prove that the SEN assumption holds, we resort to the CAV assumptions from Sennott (2009), which are sufficient for the SEN assumptions to hold. We begin with definitions that are used in the CAV assumptions.

Let $\mathcal{G}$ refer to a set of states, and $z$ refers to a single state in the final 2 definitions.

**Definition 2.** $c_{i \mathcal{G}}(\theta)$ the expected cost of a first passage from state $i$ to $\mathcal{G}$ (Sennott p. 139).

**Definition 3.** $m_{i \mathcal{G}}$ the expected time of a first passage from state $i$ to $\mathcal{G}$ (Sennott p. 140).

**Definition 4.** $R(i, \mathcal{G})$ is the set of policies, $\theta$, satisfying $\mathbb{P}_\theta(X_n \in \mathcal{G} \text{ for some } n \geq 1 | X_0 = i) = 1$ and $m_{i \mathcal{G}}$ is finite (Sennott p. 139).

$R^*(i, \mathcal{G})$ is the set of policies $\theta \in R(i, \mathcal{G})$ such that $c_{i \mathcal{G}}(\theta)$ is finite (Sennott p. 140).

**Definition 5.** Let $d$ be a stationary policy. Then $d$ is a $z$-standard policy if the Markov Chain (MC) induced by $d$ is $z$-standard.

A MC is $z$-standard if there exists a distinguished state $z$ such that $m_{iz} < \infty$ and $c_{iz} < \infty \forall i \in S$.

**Definition 6 (CAV Assumptions).**

(CAV1) There exists a $z$-standard policy $d$ with positive recurrent class $\mathcal{R}_d$.

(CAV2) Let $C(i, a)$ be the one stage cost associated with being in state $i$ and taking action $a$. Given $U > 0$, the set $D_U = \{i | C(i, a) \leq U \text{ for some } a\}$ is finite.

(CAV3) Given state $i \in \mathcal{S} / \mathcal{R}_d$, there exists a policy $\theta_i \in R^*(z, \{i\})$, where $\{i\}$ is a set containing only state $i$.

Before proving that the CAV assumptions hold in our setting, we first define a $z$-standard policy, $d$, such that CAV1 holds and that there exists a set of policies $\{\theta_i : i \in \mathcal{S} / \mathcal{R}_d\}$ for which CAV3 holds.
A $z$-standard policy. We define policy $d$ as discharging every patient in the hospital at each decision epoch. In CAV1 below, we will show this policy is a $z$-standard policy, where state $z = 0^{M \times (J+1)}$, i.e. the empty state.

**CAV 1.** Let $i_{m,j}$ be the $(m,j)^{th}$ element of the vector state $i$, and $S_i = \sum_m \sum_j i_{m,j}$. Let $z = 0^{M \times (J+1)}$.

We first show that policy $d$ is a $z$-standard policy. (i) The expected first passage time $m_{iz} = 1$ for any state $i \in S$ since the policy discharges all patients, the system state immediately returns to 0.

(ii) The expected cost from $i$ to $z$, $c_{iz}$, is finite since $c_{iz} = R \sum_m \sum_j r(m,j) \cdot i_{m,j} + C \cdot (S_i - N)^+ < \infty$.

Policy $d$ induces $\mathcal{R}_d = \{ i : i_{m,j} = 0 \ \forall j > 0 \}$, which is the set of states where no patient has stayed more than one day. We next show that $\mathcal{R}_d$ is positive recurrent under policy $d$ by showing that all states in $\mathcal{R}_d$ have a finite mean recurrence time. Consider the fact that we begin each decision epoch in the empty state $z$ (after the discharges from previous epoch). Then the mean recurrence time for any state $i \in \mathcal{R}_d$ is simply the number of trials for the arrival process $A$ to generate state $i$. Recall that, for $i \in \mathcal{R}_d$, the total number of patients that arrive each day, $S_i = \sum_m i_{m,0}$, follows a Poisson process with rate $\Lambda$ (including the readmissions), and given the total number $S_i$, the number of arrivals to each class follows multinomial distribution with parameters $\vec{p} = (p_1, p_2, \ldots, p_M)$, where $p_m$ is the probability that an arrival is of class $m$. Then

$$q_i = \mathbb{P}(A = i) = \mathbb{P}(\text{Poiss}(\Lambda) = S_i) \cdot \mathbb{P}(\text{Multinom}(S_i, \vec{p}) = i) > 0.$$  

The Poisson probability for any state is strictly positive, and so is the multinomial distribution. Thus, $q_i$ is strictly positive, and the expected number of trials to reach state $i$, which follows a geometric distribution with success probability $q_i$, is finite.

**CAV 2.**

Consider the action, $a$ that discharges no patients. Take any $U > 0$. The set $D_U$ contains at most those states $i$ s.t. $C \cdot (S_i - N) < U$, which implies that the set contains at most those states such that $S_i < (U - C)/N$, which is clearly finite.

**CAV 3.**

$S/\mathcal{R}_d = \{ i : i_{m,j} = 0 \ \forall j = 0 \}$. Consider a state $i \in S/\mathcal{R}_d$, and define policy $\theta_i$ as follows. Let $A_1, A_2, \ldots, A_J$ be the arrivals in subsequent periods. Define $i_j = (i_{1,j}, i_{2,j}, \ldots, i_{M,j})$ as the vector of all patients in state $i$ who have LOS $j$. $\theta_i$ will discharge all patients unless $A_1$ generates $i_j$. If the realization of $A_1$ is $i_j$, then $\theta_i$ discharges zero patients. Then in the next epoch, we have the starting state as $(0, i_j, 0, 0, \ldots, 0)$. In this state, the policy $\theta_i$ discharges all patients unless $A_2$ generates $i_{j-1}$. If the realization of $A_2$ is $i_{j-1}$ then we discharge no patients. Thus the starting state in the next period will be $(0, i_{j-1}, i_j, 0, 0, \ldots, 0)$. We continue this construction approach for $A_3, \ldots, A_J$, at which point we have reached state $i = (i_1, i_2, \ldots, i_J)$. Since this is a series of trials, just as in CAV 1, we again have a Geometric distribution for the time to reach state $i$ and the success probability is again positive by the same argument used in CAV 1. Hence the expected passage time to reach state $i$, $m_{z,i}$, is finite. Further, the cost in each stage is clearly finite since in each stage the policy either
discharges a finite number of patients or keep a finite number of patients in the hospital. Since the one stage cost is finite and the expected number of stages is finite, therefore the total expected cost to reach state \( i \) from state \( z \), \( c_{z(i)}(\theta) \) must be finite. □

B.2. Proof of Proposition 1

Proof. For patient type \((m, t)\), we have

\[
c_d(D_k + e_{(m, t)}) - c_d(D_k' + e_{(m, t+k'-k)}) = c_d(D_k) + c_d(e_{(m, t)}) - c_d(D_k') - c_d(e_{(m, t+k'-k)}).
\]

This equality follows by the linearity of \( c_d \) of the form (13) and the condition of \( R_e = R \) for all \( t \).

Note we can write \( c_d(e_{(m, t+k'-k)}) \) in terms of the integral of the slope of \( r(\cdot, \cdot) \)

\[
c_d(e_{(m, t+k'-k)}) = c_d(e_{(m, t)}) + R \cdot \int_t^{t+k'-k} \frac{\partial r(m; s)}{\partial s} ds.
\]

Therefore

\[
c_d(e_{(m, t)}) - c_d(e_{(m, t+k'-k)}) = -R \cdot \int_t^{t+k'-k} \frac{\partial r(m; s)}{\partial s} ds.
\]

The main result follows directly as

\[
c_d(D_k + e_{(m_1, t_1)}) - c_d(D_k' + e_{(m_1, t_1+k'-k)}) = c_d(D_k) - c_d(D_k') - R \cdot \int_{t_1}^{t_1+k'-k} \frac{\partial r(m_1; s)}{\partial s} ds \leq c_d(D_k) - c_d(D_k') - R \cdot \int_{t_2}^{t_2+k'-k} \frac{\partial r(m_2; s)}{\partial s} ds = c_d(D_k + e_{(m_2, t_2)}) - c_d(D_k' + e_{(m_2, t_2+k'-k)}).
\]

The first and last equalities follow from (42). The inequality follows from Equation (17) in the strong dominance definition.

B.3. Proof of Theorem 2

Proof. In the main paper, we have proved the case where \( t_1 + k' - k \leq J \). Now consider the case where \( t_1 + k' - k > J \), which means we cannot directly interchange type \((m_1, t_1)\) and type \((m_2, t_2)\). Let \( k'' \) represent the epoch in which the type \((m_2, t_2)\) patient has reached their maximum LOS and must be discharged. Then let policy \( \pi'_2 \) be the policy that interchanges the patients as before except discharges the type \((m_2, t_2)\) patient at \( k'' \) instead of \( k' \). If \( V^{\pi'_2} - V^{\pi_1} \leq 0 \) then the argument still works. Otherwise we have that \( V^{\pi'_2} > V^{\pi_1} \). Since the readmission costs are higher in \( V^{\pi'_2} \) than in \( V^{\pi_2} \) — the only difference between the two policies is that one patient is discharged earlier, which incurs a higher discharge cost — it means that the savings in terms of holding cost by discharging the patient earlier, which we call \( \Delta c_h \), must be greater than the increase in readmission cost, \( c_d(D_k'' + e_{(m_2, t_2+k''-k)}) - c_d(D_k' + e_{(m_2, t_2+k'-k)}) \). We now construct a corresponding policy for \( \pi_1 \), which we call \( \pi'_1 \) and show that it has a lower cost than \( \pi_1 \) and hence we can substitute \( \pi'_1 \) for \( \pi_1 \) in the interchange argument. That is, let \( \pi'_1 \) be the policy that discharges the single type \((m_1, t_1)\) patient at \( k'' \) instead of \( k' \). Then

\[
V^{\pi'_1} - V^{\pi_1} = [c_d(D_k'' + e_{(m_1, t_1+k''-k)}) - c_d(D_k' + e_{(m_1, t_1+k'-k)})] - \Delta c_h 
\leq c_d(D_k'' + e_{(m_2, t_2+k''-k)}) - c_d(D_k' + e_{(m_2, t_2+k'-k)}) - \Delta c_h \leq 0
\]
The first equality holds because we are only changing the discharge timing of the target patient, which incurs a readmission penalty of \( c_d(D_{k'} + e_{(m_1,t_1+k''-k)}) - c_d(D_{k'} + e_{(m_1,t_1+k'-k)}) \) since the patient is discharged earlier, but gains the same occupancy benefit of policy \( \pi_2' \), i.e. \( \Delta ch \), since the occupancy is the same for both policies \( \pi_1' \) and \( \pi_2' \). The first inequality follows from Proposition 1, which implies that \( c_d(D_{k'} + e_{(m_1,t_1+k''-k)}) - c_d(D_{k'} + e_{(m_1,t_1+k'-k)}) \leq c_d(D_{k'} + e_{(m_2,t_2+k''-k)}) - c_d(D_{k'} + e_{(m_2,t_2+k'-k)}) \). The final inequality follows because that is the assumption above that the occupancy cost reduction outweighs the readmission cost increase, which is the only scenario in which the direct interchange argument does not work. Since \( \pi_1' \) is a better policy than \( \pi_1 \), we can replace \( \pi_1 \) with \( \pi_1' \) and once again make the exact same direct interchange argument as before except using \( \pi_1' \) and \( \pi_2' \). That is \( V_{\pi_1} - V_{\pi_1} \leq V_{\pi_2} - V_{\pi_1} \leq 0 \), where the last inequality follows by using the interchange argument above. \( \square \)

B.4. Proof of Proposition 3

Proof. We use an induction argument. We start from proving the proposition in period \( T \), and then prove for general period \( k \). Let \( S_A = E(1 \cdot A_k) = E(\sum_m A_k^m) \) and \( S_A^2 = E(1 \cdot A_k)^2 = E(\sum_m A_k^m)^2 \) be the first and second moment of the total number of arrivals on each day.

**Period \( T \).** On period \( T \), Thus, the value function on period \( T \) is given by

\[
V_T(X_T) = \min_{D_T^j, j=0,...,J} C(S_T)^2 + \sum_j D_T^j R_j D_T^j + C(S_T - \sum_j 1 \cdot D_T^j)^2 + 2CS_A(S_T - \sum_j 1 \cdot D_T^j) + CS_A^2.
\]

The terms on the last row come from the fact that

\[
E(S_{T+1})^2 = E_{A_T}(S_T - \sum_j 1 \cdot D_T^j + A_T)^2
\]

when we evaluate the cost-to-go function \( EV_{T+1}(X_{T+1}) \).

For notational simplicity, we focus on the single-class setting with \( M = 1 \) in the rest of the proof; the derivation for multi-class setting is similar using matrix calculus. Taking the derivative of \( V_T(X_T) \) with respect to each \( D_T^j \), we have that

\[
\frac{\partial V_T(X_T)}{\partial D_T^j} = 2R_j D_T^j - 2C(S_T - \sum_j 1 \cdot D_T^j) - 2CS_A, \quad j = 0,\ldots,J.
\]

Setting these derivatives to be 0, we get the optimal number of discharges \( D_T^* \) satisfying

\[
D_T^* = (D_T^{0*}, D_T^{1*}, \ldots, D_T^{j*})' = U_T^{-1}S_T + U_T^{-1}B_T,
\]

where

\[
U_T = \begin{pmatrix}
C + R_0 & C & \cdots & C \\
C & C + R_1 & \cdots & C \\
& \ddots & \ddots & \ddots \\
C & C & \cdots & C + R_J
\end{pmatrix}, \quad B_T = \begin{pmatrix}
CS_A \\
CS_A \\
& \ddots \\
CS_A
\end{pmatrix}, \quad S_T = \begin{pmatrix}
CS_A \\
CS_T \\
& \ddots \\
CS_T
\end{pmatrix}.
\]

Author: Article Short Title
Note that the number of discharges is linear in the total number of patients in system, \( S_T \). Let \( a^j_T \) be the sum of entries on row \( j \) of matrix \( C U_T^{-1} \), and \( b^j_T \) denote the \( j \)th row of \( U_T^{-1} B_T \). We can further write each optimal discharge quantity \( D^j_T \) as

\[
D^j_T = a^j_T S_T + b^j_T, \quad j = 0, \ldots, J,
\]

where \( b^j_T \) is independent of the system state \( X_T \).

Plugging the (unconstraint) optimal discharge quantity \( D^j_T \) to \( V_T(X_T) \) and after some algebra,

\[
V_T(X_T) = CS_T^2 + \sum_j R_j (a^j_T S_T + b^j_T)^2 \\
+ C (S_T - \sum_j (a^j_T S_T + b^j_T))^2 + 2 C S_A (S_T - \sum_j (a^j_T S_T + b^j_T)) + C S_A^2 \\
= \alpha_T S_T^2 + \theta_T S_T + \kappa_T,
\]

where

\[
\alpha_T = C + \sum_j R_j (a^j_T)^2 + C (1 - \sum_j a^j_T)^2, \\
\theta_T = 2 \sum_j R_j a^j_T b^j_T - 2 C (1 - \sum_j a^j_T) \sum_j b^j_T + 2 C S_A (1 - \sum_j a^j_T),
\]

and

\[
\kappa_T = \sum_j R_j (b^j_T)^2 + C (\sum_j b^j_T)^2 - 2 C S_A \sum_j b^j_T + C S_A^2,
\]

with the last term \( \kappa \) being independent of \( X_T \).

**Period \( k \).** Assume period \( k + 1 \) has been proved. The value function on period \( k \) equals

\[
V_k(X_k) = \min_{D^j_k = 0, \ldots, J} CS_k^2 + \sum_j R_j (D^j_k)^2 \\
+ \alpha_{k+1} E(S_k - \sum_j D^j_k + A_k)^2 + \theta_{k+1} E(S_k - \sum_j D^j_k + A_k) + \kappa_{k+1}
\]

\[
= \min_{D^j_k = 0, \ldots, J} CS_k^2 + \sum_j R_j (D^j_k)^2 \\
+ \alpha_{k+1} (S_k - \sum_j D^j_k)^2 + 2 \alpha_{k+1} S_A (S_k - \sum_j D^j_k) + \alpha_{k+1} S_A^2 + \theta_{k+1} (S_k - \sum_j D^j_k) + \theta_{k+1} S_A + \kappa_{k+1}.
\]

Taking the derivative of \( V_k(X_k) \) with respect to each \( D^j_k \), we get

\[
\frac{\partial V_k(X_k)}{\partial D^j_k} = 2 R_j D^j_k - 2 \alpha_{k+1} (S_k - \sum_j D^j_k) - 2 \alpha_{k+1} S_A - \theta_{k+1}, \quad j = 0, \ldots, J.
\]

Setting these derivatives to be 0, we get the optimal number of discharges \( D^j_k \) satisfying

\[
D^j_k = (D_0^j, D_1^j, \ldots, D^j_{J})' = U_k^{-1} S_k + U_k^{-1} B_k,
\]

where

\[
U_k = \begin{pmatrix} 
\alpha_{k+1} + R_0 & \alpha_{k+1} & \ldots & \alpha_{k+1} \\
\alpha_{k+1} & \alpha_{k+1} + R_1 & \ldots & \alpha_{k+1} \\
\ldots & \ldots & \ldots & \ldots \\
\alpha_{k+1} & \alpha_{k+1} & \ldots & \alpha_{k+1} + R_J 
\end{pmatrix}, \quad B_k = \begin{pmatrix} 
\alpha_{k+1} S_A + \theta_{k+1}/2 \\
\alpha_{k+1} S_A + \theta_{k+1}/2 \\
\ldots \\
\alpha_{k+1} S_A + \theta_{k+1}/2 
\end{pmatrix}, \quad S_k = \begin{pmatrix} 
\alpha_{k+1} S_k \\
\alpha_{k+1} S_k \\
\ldots \\
\alpha_{k+1} S_k 
\end{pmatrix}.
\]
Equivalently, we can rewrite $D_k^*$ as

$$D_k^* = a_k^i S_k + b_k^i, \quad j = 0, \ldots, J,$$

where $b_j$ is independent of the system state $X_k$. Plugging back to $V_k$, we get

$$V_k(X_k) = CS_k^2 + \sum_j R_j(a_k^i S_k + b_k^i)^2$$

$$+ \alpha_{k+1}(S_k - \sum_j a_k^i S_k - \sum_j b_k^i)^2 + 2\alpha_{k+1} S_A(S_k - \sum_j a_k^i S_k - \sum_j b_k^i) + \alpha_{k+1} S_A^2$$

$$+ \theta_{k+1}(S_k - \sum_j a_k^i S_k - \sum_j b_k^i) + \theta_{k+1} S_A + \kappa_{k+1}$$

$$= \alpha_k S_k^2 + \theta_k S_k + \kappa,$$

where

$$\alpha_k = C + \sum_j R_j(a_k^i)^2 + \alpha_{k+1}(1 - \sum_j a_k^i)^2,$$

$$\theta_k = 2\sum_j R_j a_k^i b_k^i - 2\alpha_{k+1}(1 - \sum_j a_k^i) \sum_j b_k^i + 2(\alpha_{k+1} S_A + \theta_{k+1}/2)(1 - \sum_j a_k^i),$$

and

$$\kappa_k = \sum_j R_j(b_k^i)^2 + \alpha_{k+1}(\sum_j b_k^i)^2 - 2(\alpha_{k+1} S_A + \theta_{k+1}/2) \sum_j b_k^i + \alpha_{k+1} S_A^2 + \theta_{k+1} S_A + \kappa_{k+1},$$

with the last term $\kappa$ being independent of $X_k$. □

**B.5. Proof for convexity of the predicted readmission function**

We now show that the predicted readmission probability function, $\bar{S}$ is a convex function in LOS for each patient. Fix a patient $j$. By plugging (1) and (2), we can write (3) as

$$\bar{S}(t^*; Y_j) = \frac{1}{1 + e^{Y_j^T A}_j} \sum_{w=1}^W \pi_w \left(1 - \exp \left(- e^{Y_j^T \beta_w} \int_0^{t^*} h_0^w(u) du \right) \right).$$

Recall that $Y_j = (\log(LOS_j), Y_j^c)$. To evaluate the (marginal) impact of changing $LOS_j$ on her $t^*$-day readmission probability, we fix $Y_j^c$ and let $\xi = (\xi_{\ell}, \xi_c)$ and $\beta_w = (\beta_{\ell w}, \beta_c^w)$ be the coefficients associated with $LOS_j$ and $Y_j^c$, respectively. Using the fact that $e^{a \log b} = b^a$, we can write $\bar{S}$ as a function of $LOS_j$:

$$\bar{S}(t^*, LOS_j; Y_j^c) = \frac{1}{1 + (LOS_j)^{\xi_{\ell}} A_j} \sum_{w=1}^W \pi_w \left(1 - \exp \left(- (LOS_j)^{\beta_{\ell w}} B_j \right) \right),$$

where

$$A_j = e^{Y_j^T \xi_c}, \quad B_j = e^{Y_j^T \beta_c} \int_0^{t^*} h_0^w(u) du,$$

which are two terms not depending on $LOS_j$. Next, we state the lemma on the convexity of $\bar{S}$.

**Lemma 1.** When $\xi_{\ell} > 0$ and $\beta_{\ell w} < 0$ for $w = 1, \ldots, W$, the function $\bar{S}$ is convex in $LOS$ for given $t^* > 0$ and $Y_j^c$. 

Proof. It is sufficient to prove $S(x)$ is convex in $x = \text{LOS}_j$. To see this, let $S_1(x) = \frac{1}{1 + x^{\beta_{\ell}} \cdot A_j}$ and $S_2(x) = \sum_{w=1}^{W} \pi_w \left(1 - \exp \left(-x^{\beta_{\ell}} \cdot B_j \right) \right)$. Note that LOS, $A_j$ and $B_j$ are always positive, thus, it is easy to prove that both $S_1(x)$ and $S_2(x)$ are convex in $x$ when the conditions for $\xi_{\ell}$ and $\beta_{w}^{w}$ are satisfied. Now we prove the product of $S_1$ and $S_2$ is convex. We have

$$S_1(x_1)S_2(x_1) - S_1(x_2)S_2(x_2) = S_1(x_1)(S_2(x_1) - S_2(x_2)) + S_2(x_2)(S_1(x_1) - S_1(x_2)).$$

Similarly, we have

$$S_1(x_2)S_2(x_2) - S_1(x_3)S_2(x_3) = S_1(x_2)(S_2(x_2) - S_2(x_3)) + S_2(x_3)(S_1(x_2) - S_1(x_3)).$$

Let $x_0 = \frac{x_1 + x_3}{2}$. Due to the convexity of $S_1$ and $S_2$, we have $S_1(x_1) - S_2(x_2) \geq (S_2(x_2) - S_2(x_3))$ and $(S_1(x_1) - S_1(x_3)) \geq S_1(x_2) - S_1(x_3)$). Since all $S_1$ and $S_2$ are positive numbers, we have $S_1(x_1)S_2(x_1) - S_1(x_2)S_2(x_2) \geq S_1(x_2)S_2(x_2) - S_1(x_3)S_2(x_3)$. Using the fact that $S(x)$ is a continuous function, midpoint convexity implies convexity and the proof is complete. $\square$

B.6. Static threshold optimization model

Denote the set of thresholds as $(l_1, \ldots, l_M)$ under the static threshold setting. Let the action be, $D^m_k$, the number of class $m$ discharges on day $k$. To construct a threshold-based discharge policy, we restrict the action space in Equation (15) to be $\Pi_s = \{ D^m_k : D^m_k = \sum_{j=l_m}^{J} X^m_{k,j}, \forall X_k \in \mathbb{N}^{M \times (J+1)}, k = 1, 2, \ldots, m = 1, \ldots, M \} \subset \Pi$. Note, after one round of discharge there will never be a patient of class $m$ with LOS longer than $l_m$, thus, all patients being discharged are of LOS $l_m$. That is,

$$D^m_k = X^m_{k,l_m}$$

in steady-state. Letting $Q_k$ be the pre-action queue length on day $k$ under the policy $\pi \in \Pi_s$, the average cost objective function is given by:

$$Z^\pi = \limsup_{T \to \infty} \frac{1}{T} \sum_{k=0}^{T} \sum_{m=1}^{M} R \cdot r(m, l_m) \mathbb{E}[D^m_k] + C \cdot \mathbb{E}[Q_k]$$

$$Z^* = \inf_{\pi \in \Pi_s} Z^\pi$$

Through an easy modification of the proof Theorem 1, we can show that there exists an optimal stationary policy for objective function in (46).

Proof of Proposition 2. The external arrival rate (excluding readmission) is $\Lambda$. In the single-class setting, solving the flow balance equation (20) gives us the total arrival rate (including readmissions) as

$$\tilde{\Lambda} = \Lambda^c / (1 - r(\ell)).$$

For a given $\ell$, the offered load on the system is

$$B = B(\ell) = \ell \cdot \tilde{\Lambda} = \Lambda \cdot \frac{\ell}{1 - r(\ell)}.$$
Using $B$, the objective function $f(\ell)$ (based on (46)) can be written as

$$f(\ell) = R\tilde{\Lambda} \cdot r(\ell) + C \cdot \mathbb{E}[\text{Poiss}(B) - N]^+$$. \hfill(47)

For $\mathbb{E}[\text{Poiss}(B) - N]^+$, when $N$ and $B$ are reasonably large, we can approximate $\text{Poiss}(B)$ with a normal random variable with mean $\mu = B$ and standard deviation $\sigma = \sqrt{B}$. Let $\alpha = \frac{N - B}{\sqrt{B}}$. With the Normal approximation assumption made, the mean queue length can be approximated by

$$\mathbb{E}[\text{Poiss}(B) - N]^+ \approx \mathbb{E}\left[\text{Norm}(B, \sqrt{B}) - N\right]^+ = \sqrt{B} [\phi(\alpha) - \alpha(1 - \Phi(\alpha))] = \sqrt{B}\phi(\alpha) - (N - B)(1 - \Phi(\alpha)).$$

Taking its derivative with respect to $B$, we get

$$\frac{\phi(\alpha)}{2\sqrt{B}} + \sqrt{B}\phi'(\alpha)\alpha'(B) + (N - B)\phi(\alpha)\alpha'(B) + (1 - \Phi(\alpha)).$$

When $B$ is close to $N$, which we expect in most congested hospitals,

$$\alpha'(B) = -\frac{N + B}{2B\sqrt{B}} \approx -\frac{1}{\sqrt{B}}, \quad \phi'(\alpha) = -\alpha\phi(\alpha).$$

Also, when $B$ is large, $\frac{\phi(\alpha)}{2\sqrt{B}} \approx 0$. Thus, we can further simplify the previous derivative as

$$\frac{d\mathbb{E}[\text{Poiss}(B) - N]^+}{dB} \approx \alpha\phi(\alpha) - \frac{N - B}{\sqrt{B}}\phi(\alpha) + (1 - \Phi(\alpha)) = (1 - \Phi(\alpha)).$$

Next, we take the derivative of $f(\ell)$ in (47) with respect to $\ell$ and set it to 0 to find the first order condition on the optimal discharge threshold. We then have

$$R \cdot \left(\frac{d\Lambda}{d\ell} r(\ell) + r'(\ell)\Lambda\right) = C\left(1 - \Phi(\alpha)\right) \frac{d}{d\ell} B.$$ \hfill(48)

□

**Appendix C: A decomposition heuristic for weak dominance ranking**

The strong dominance ranking criterion we proved in Section 5.1 of the main paper requires dominance in the reduction in readmission risk between the current day and any future day between patient classes. Although our empirical study shows this strong dominance holds among patients in our dataset, it can be violated in a different hospital. In this section, we develop a decomposition heuristic that allows us to specify a useful weak dominance ranking criterion that is more general than the strong dominance criterion. This decomposition method leverages the linear quadratic approximation developed in Section 6.1 and focuses on a single patient type (class $m$ and LOS $j$). It helps identify determinants that impact how many patients of each type should be discharged in the current environment. Factors that drive higher discharge rates are then used to rank patients in terms of “desirability” of discharging that patient type. In the end, we can see that this weak dominance ranking criterion considers the entire trajectory of each patient class, $m$, while it covers the strong dominance criterion as a special case.
Heuristic 1 Patient type decomposition

1: for $m \in 1, \ldots, M$ do
2: Select class $m$. Do backward Induction on LOS.
3: for $\ell = J - 1, J - 2, \ldots, 1$ do
4: Objective: optimize $D^m,\ell_k$
5: Step 1: Day $k$ Actions
6: for all $(n,j) \neq (m,\ell)$ do
7: $D^{n,j}_{k} \leftarrow a^{n,j}_{k} S_{k} + b^{n,j}_{k}$ \hspace{1cm} \triangleright \text{Linear quadratic solution for all other patients}
8: end for
9: Step 2: Day $k + 1, \ldots, k + \ell$ Actions
10: for $t = 1, \ldots, J - \ell$ do
11: if $t = J - \ell$ then
12: $D^{m,J}_{k+t} = X^{m,J}_{k+t} - \sum_{t'=0}^{J-\ell} D^{m,J-t}_{k+t}$ \hspace{1cm} \triangleright \text{Discharge all patients reaching max LOS}
13: else
14: for all $(n,j) \neq (m,J-t)$ do
15: $S_{k+t} \leftarrow S_{k+t-1} + A_{k+t-1} - \sum_{n',j'} D^{n',j'}_{k+t-1}$
16: $D^{n,j}_{k+t} \leftarrow a^{n,j}_{k+t} S_{k+t} + b^{n,j}_{k+t}$ \hspace{1cm} \triangleright \text{Linear quadratic solution for all other patients}
17: end for
18: $X^{n,j+1}_{k+t+1} \leftarrow X^{n,j}_{k+t} - D^{n,j}_{k+t}$
19: end if
20: end for
21: $D^m,\ell_k$ Optimal Cost Criterion
22: $S_{k+\ell+1} \leftarrow S_{k+\ell}(D^m,\ell_k) + A_{k+\ell} - D^{m,J}_{k+\ell} - \sum_{(n,j) \neq (m,J)} D^{n,j}_{k+\ell}$
23: $Z^{\pi}_{dep} \leftarrow \sum_{t=0}^{J-\ell} [c_h(X_{k+t}) + c_d(D_{k+t})] + V^*_{LQ}(S_{k+\ell+1}(D^m,\ell_k))$ \hspace{1cm} \triangleright \text{Note: $D_{k+t}$ and $X_{k+t}$ are vectors including all patient types.}
24: Structure of optimal discharge action
25: Solve optimal $D^m,\ell_k$ from FOC $\frac{\partial Z^{\pi}_{dep}}{\partial D^m,\ell_k} = 0$.
26: end for
27: end for
28: Output:
29: Weak dominance rank in Equation (49)

C.1. Patient Type Decomposition

The recursive approach we take to determining the weak dominance ranking and the univariate transformation is detailed in Heuristic 1 below.

To explain, for a given class $m$, We first focus on patients who have stayed $J - 1$ days in the hospital in the current period, $k$. We individually optimize the discharge action, $D^m,J-1_k$, for these patients,
type \((m, J - 1)\), in period \(k\) as follows. We let the optimal actions of all other patients in period \(k\) be determined by the linear quadratic approximation given in Equation (23) of Proposition 3. In period, \(k + 1\), we know that the the patients of type \((m, J - 1)\) who are not discharged in period \(k\), i.e. \(X^{m,J-1}_k - D^{m,J-1}_k\), will be discharged because they have reached their maximum LOS. For the other types of patients, we again use the linear quadratic solution for the optimal response actions. Finally, we approximate the cost-to-go by \(Z^*_LQ\) where the solution is given in (24) of Proposition 3.

With this decomposition approach, we can optimize \(D^{m,J-1}_k\) using the first order condition.

Next, we consider class \(m\) patients who have stayed \(J - 2\) days in the hospital in the current period, \(k\). Again, we individually optimize the discharge action, \(D^{m,J-2}_k\) in the same manner as described in the previous paragraph. Two major differences here: (i) \(D^{m,J-2}_k\) will not only impact the current and subsequent periods, \(k\) and \(k + 1\), but will also impact period \(k + 2\), since there are now two days until type \((m, J - 2)\) patients to reach their maximum LOS; (ii) the patients who are not discharged, \(X^{m,J-2}_k - D^{m,J-2}_k\), will become type \((m, J - 1)\) on day \(k + 1\), and for day \(k + 1\)’s discharge decision on type \((m, J - 1)\), \(D^{m,J-2}_{k+1}\), we will use the optimal discharge action obtained in the previous paragraph.

We demonstrate at the end of this section that the optimal action for \(D^{m,J-2}_{k+1}\) is again a linear function in \(X^{m,J-2}_k\) and \(D^{m,J-2}_k\). Actions for other patient types in period \(k + 1\) and \(k + 2\) are approximated again by the linear decision rules, which will depend on \(D^{m,J-2}_k\).

We then recursively work backward to solve for the optimal discharge action for the remainder of the possible Lengths of Stay, \(J - 3, J - 4, \ldots, 1\). Eventually, from the structure of the optimal actions resulting from Heuristic 1, we are able to identify the main drivers for the discharge quantity, which motivate the following ranking heuristic:

\[
\psi^{m,j}_1(R_{m,j} - R_{m,j+1}) + \psi^{m,j}_2(R_{m,j} - R_{m,j+2}) + \cdots + \psi^{m,j}_{J-1}(R_{m,j} - R_{m,j}),
\]

where \(R_{m,j} = R \cdot r(m,j)\) for type \((m,j)\) patients (also see Equation (27) in the main paper). This ranking heuristic is a weighted average based on the marginal changes across the entire remaining risk trajectory for different patient classes. We call this ranking criterion weak dominance. In the extreme case where all \(\psi_i\)’s are 0, the score is simply \(R_{m,j} - R_{m,j}\); if all \(\psi_i\)’s are 0 except \(\psi_{k-1} = -1\), the score is \(R_{m,j} - R_{m,j+1}\).

Similar as the strong dominance ranking, a lower score from (49) indicates a higher discharge number, as the patient gains lower marginal benefit from staying in the hospital longer; i.e., it is unlikely that they will be greatly improved by extending their stay. We will demonstrate the performance of this weak dominance ranking criterion via numerical experiments in Appendix D.1.

**Remark 8.** Strong dominance from the structural properties implies a weak dominance.

### C.2. Derivation details

In this section, we specify more details on getting the optimal action for \(D^{m,J-1}_k\) and \(D^{m,J-2}_k\) in the decomposition heuristic; the actions for other types can be derived similarly. Throughout this section, we use \(R_{m,j} = R \cdot r(m,j)\) to denote the expected readmission cost for type \((m,j)\) patients.
LOS $J - 1$ on day $k$, class $m$.

The current period cost is $CS_k^2 + R_{m,J-1}D_{k}^{m,J-1} + \sum_{n,j:(n,j)\neq (m,J-1)} R_{n,j}(a_k^{n,j}S_k + b_k^{n,j})$. Note that we assume all discharge actions are going to follow the linear decision rule, except discharge action $D_{k}^{m,J-1}$ for the group of patients we trace. On day $k+1$, we know that all the remaining patients $X_k^{m,J-1} - D_{k}^{m,J-1}$ reach LOS of $J$ days and have to be discharged. Thus, period $k+1$ expected cost and future cost-to-go is

$$E \left[ CS_{k+1}^2 + R_{m,J}(X_k^{m,J-1} - D_{k}^{m,J-1}) + \sum_{n,j:(n,j)\neq (m,J)} R_{n,j}(a_k^{n,j}S_{k+1} + b_k^{n,j}) + V_{k+2}(S_{k+2}) \right],$$

where $V_{k+2}(S_{k+2}) = \alpha_{k+2}S_{k+2}^2 + \theta_{k+2}S_{k+2} + \kappa_{k+2}$ comes from the linear quadratic approximation.

Taking the derivative of period $k$, $k+1$ cost and cost-to-go with respect to action $D_{k}^{m,J-1}$ and after some algebra, we have the first-order condition as

$$(R_{m,J-1} - R_{m,J}) + \sum_{n,j:(n,j)\neq (m,J)} R_{n,j}a_k^{n,j} = 2C(\bar{S}_{k+1} + E[A_k]) + 2(\alpha_{k+2}(\bar{S}_{k+2} + E[A_k + A_{k+1}]) + \beta_{k+2}/2) \sum_{n,j:(n,j)\neq (m,J)} a_k^{n,j}, \quad (50)$$

where $\bar{S}_{k+1}$ and $\bar{S}_{k+2}$ depend on $D_{k}^{m,J-1}$ via

$$\bar{S}_{k+1} = S_k - D_{k}^{m,J-1} - \sum_{n,j:(n,j)\neq (m,J-1)} (a_k^{n,j}S_k + b_k^{n,j}),$$

and

$$\bar{S}_{k+2} = \left(1 - \sum_{n,j:(n,j)\neq (m,J)} a_k^{n,j}\right) \bar{S}_{k+1} - (X_k^{m,J-1} - D_{k}^{m,J-1}) - \sum_{n,j:(n,j)\neq (m,J)} b_k^{n,j}.$$

Note that these two partial sums are constants when we fix an $S_k$ and $D_{k}^{m,J-1}$. In Equation (50), on the left-hand side, the first term $(R_{m,J-1} - R_{m,J})$ shows the effect of $D_{k}^{m,J-1}$ on readmission cost reduction for the group of patients we trace, and the second term $\sum_{n=1}^{M} \sum_{j:(n,j)\neq (m,J)} R_{n,j}a_k^{n,j}$ reflects the approximate effect on the changes of readmission cost for all patients discharging on day $k+1$. The terms on the right-hand side reflects the gain on congestion cost reduction, with the first term reflecting the effect on day $k+1$ congestion, and the second term the effect on future costs (approximated by $V_{k+2}$).

Optimal action. Solving the first-order equation (50), we have the unconstrained optimal action as a linear function of $S_k$. To illustrate, consider a special case where $\alpha_{k+2} = \beta_{k+2} = 0$, then

$$D_{k}^{m,J-1} = S_k \left(1 - \sum_{n,j:(n,j)\neq (m,J-1)} a_k^{n,j}\right) - \sum_{n,j:(n,j)\neq (m,J)} b_k^{n,j} + E[A_k] - \frac{(R_{m,J-1} - R_{m,J}) + \sum_{n,j:(n,j)\neq (m,J)} R_{n,j}a_k^{n,j}}{2C}, \quad (51)$$

In other words, we can write $D_{k}^{m,J-1}$ as a linear function of $S_k$:

$$D_{k}^{m,J-1} = \psi_{m,J-1}S_k + \phi.$$

(52)
In the above special case example, we have \( \hat{\psi}^{m,j-1} = \left(1 - \sum_{n,j \neq (m,j)} a^n_{m,j} \right) \).

**Ranking.** Note that \( \sum_{k=1}^{M} \sum_{(n,j) \neq (m,j)} R_{n,j} \alpha^{n,j}_{m,j} \) are roughly the same when the readmission penalty cost \( R \) is reasonably large comparing to \( C \) (\( R \geq C \) as supported from our numerical experiments). Then the dominating term in Equation (50) is \( (R_{m,j-1} - R_{m,j}) \). The larger this difference is, the smaller amount we discharge. Thus, we should give priority to discharge patients with a smaller \( (R_{m,j-1} - R_{m,j}) \).

Note that consistent with the linear-quadratic approximation, the optimal action is still linear in \( S_k \). The main change is that the ranking is based on \( R_{m,j-1} - R_{m,j} \), instead of \( R_{m,j-1} \). This is because we explicitly consider that all patients have to be discharged when reaching a LOS of \( J \) days.

**LOS \( J - 2 \) on day \( k \), class \( m \).**

Next, we look at type \((m,j-2)\) patients. Again, we assume all discharge actions follow the linear actions, except \( D_{k,j}^{J-2} \) and \( D_{k+1,j}^{J-1} \). The remaining patients \( X_{k,j}^{J-2} - D_{k,j}^{J-2} - D_{k+1,j}^{J-1} \) who are not discharged on day \( k \) or \( k+1 \) will all be discharged on day \( k+2 \) as reaching LOS of \( J \) days.

The cost on day \( k \) is \( CS_k^2 + R_{m,j-2} D_{k,j}^{m,j-2} + \sum_{n,j \neq (m,j-2)} R_{n,j} (a^n_{m,j} S_k + b^n_{m,j}) \). Period \( k+1 \) and \( k+2 \) expected cost and cost-to-go function equal

\[
E \left[ CS_{k+1}^2 + R_{m,j-1} D_{k+1,j}^{m,j-1} + \sum_{n,j \neq (m,j-1)} R_{n,j} (a^n_{k+1,j} S_{k+1} + b^n_{k+1,j}) \right] \\
+ E \left[ CS_{k+2}^2 + R_{m,j} (X_{k+2,j}^{J-2} - D_{k+2,j}^{J-2} - D_{k+1,j}^{J-1}) + \sum_{n,j \neq (m,j)} R_{n,j} (a^n_{k+2,j} S_{k+2} + b^n_{k+2,j}) + V_{k+2} (S_{k+2}) \right].
\]

Using a similar argument as above, using the first order condition, the optimal action of \( D_{k,j}^{J-1} \) will balance the readmission cost changes of the traced patients, the readmission costs of all patients in periods \( k, k+1 \) and \( k+2 \), as well as the congestion costs in periods \( k, k+1 \) and \( k+2 \) and the cost-to-go function.

**Ranking.** To determine patient ranking, the dominant terms are readmission cost associated with \( D_{k,j}^{J-2}, D_{k+1,j}^{J-1} \), and \( X_{k,j}^{J-2} - D_{k,j}^{J-2} - D_{k+1,j}^{J-1} \). Note that the unconstrained optimal action of \( D_{k+1,j}^{m,j-1} \) linearly depend on \( S_{k+1} \) as given by (52), where \( S_{k+1} \) depends on \( D_{k,j}^{J-2} \) via

\[ S_{k+1} = S_k - D_{k,j}^{J-2} - \sum_{n,j \neq (m,j-2)} (a^n_{k,j} S_k + b^n_{k,j}) + A_k. \]

Due to the nice linear property, we can separate the expectation on the arrival quantity \( A_k \) and the other terms. This gives us the following response function of \( D_{k,j}^{m,j-1} \):

\[ D_{k,j}^{m,j-1} = -\hat{\psi}^{m,j-1} D_{k,j}^{m,j-2} + \hat{\varphi}', \]

which linearly depends on \( D_{k,j}^{m,j-2} \), with \( \hat{\psi}^{m,j-1} \) given in (52) and \( \hat{\varphi}' \) being a constant. Then, the ranking score is a linear combination between \( R_{m,j-2} - R_{m,j-1} \) and \( R_{m,j-1} - R_{m,j} \):

\[
R_{m,j-2} - R_{m,j-1} \hat{\psi}^{m,j-1} - R_{m,j} (1 - \hat{\psi}^{m,j-1}) \\
= (R_{m,j-2} - R_{m,j}) - \hat{\psi}^{m,j-1} (R_{m,j-1} - R_{m,j}) \\
= \hat{\psi}_1^{m,j-2} (R_{m,j-2} - R_{m,j-1}) + (1 - \hat{\psi}_1^{m,j-2}) (R_{m,j-2} - R_{m,j}).
\]
To interpret, the parameter $\psi_1^{m,J-2}$ reflects the probability of the patient being discharged tomorrow or two days later. Indeed, if $\psi_1^{m,J-2} = -1$, then this score simply becomes $(R_{m,J-2} - R_{m,J-1})$; if $\psi_1^{m,J-2} = 0$, then this score simply becomes $(R_{m,J-2} - R_{m,J})$. In other words, $\psi_1^{m,J-2}$ reflects how much weight we need to consider for the future risk readmission cost.

Appendix D: Additional numerical results

D.1. Performance of the dynamic algorithm: small-scale MDP

To evaluate the performance of our dynamic heuristic algorithm developed in Section 6, we test it on a small-scale MDP where the exact analysis using value iteration is still feasible. In this small-scale MDP, we consider two classes of patients with the maximum LOS $J = 3$ days. For the risk trajectory, we use two different designs: one in which strong dominance holds, and one in which it does not hold, and we instead employ weak dominance.

- Strong dominance: $r(1,j) = \{1, 0.3, 0.2, 0.15\}$; $r(2,j) = \{1, 0.3, 0.08, 0.06\}$.
- Weak dominance: $r(1,j) = \{1, 0.3, 0.21, 0.15\}$; $r(2,j) = \{1, 0.3, 0.19, 0.17\}$.

The arrival process of each patient class follows a Poisson process with the mean arrival rate being 2 patients per day. In this small-scale MDP, we also ignore the readmission arrivals for tractability of value iteration. We set $C = 1$ for the queue-length cost $c_h(X_k) = C(S_k - N)^+$, and $R = 6$ for the discharge cost $c_d(D_k) = \sum R \cdot r(m,j) D_k^{m,j}$. We vary the capacity $N = 6, 12, 20$ to experiment with high to light load conditions. We impose a cutoff for each $x^{m,j} \leq 11$ and a cutoff for each $a^{m,j} \leq 10$.

In the main paper, we have demonstrated the performance of the dynamic heuristic when the strong dominance satisfied. In the following subsection, we focus on the performance where the strong dominance is no longer satisfied and we need to use the weak dominance ranking criterion.

Modified weak dominance criterion. The weighting terms, $\psi_k^{m,j}$ in (49), account for the likelihood that the patient will be discharged on day $j + 1$ of their trajectory in the system. The more likely that the patient is to be discharged on day $j + 1$ of their trajectory, the higher the weight. Unfortunately, these weighting terms depend on the cost structures, and in certain cases, they are not easy to calculate. Motivated by their interpretation, we consider a modified ranking criterion where these weighting terms can be easily obtained. This modified ranking utilizes the static policy, where a class $m$ patient is discharged at the optimal static threshold $l_m$. In other words, $\psi_l^{m,j} = 1$ and other $\psi$’s are 0. This produces the following ranking criterion:

$$\tilde{\omega}(m,j) = R_{m,j} - R_{m,j-1}.$$  

Numerical results. Under the weak dominance setting, where $r(1,j) = \{1, 0.3, 0.21, 0.15\}$; $r(2,j) = \{1, 0.3, 0.19, 0.17\}$. Figure 14 shows that our dynamic heuristic is able to achieve almost the same (optimal) actions as those solved from value iteration. When the system is more concerned about congestion ($N = 6$), and most patients are discharged after one day, then class 1 patients should be discharged first since $(0.3 - 0.21) = 0.09 < (0.3 - 0.19) = 0.11$. However, when the system is less congested ($N = 20$) and most patients are able to stay for two days, then class 2 patients should be discharged before class 1, since $(0.3 - 0.15) = 0.15 > (0.3 - 0.17) = 0.13$. Figure 14c shows the poor performance if we use the wrong ranking.
Figure 14 Small-scale MDP with weak dominance ranking. We set $N = 6$ or 20, $C = 1$ and $R = 6$.

D.2. Comparison with other heuristics: simulation experiments

To go beyond the comparison in small-scale MDPs, we compare the dynamic heuristic algorithm with several other heuristics in a simulation environment. The experiment settings are the same as the baseline introduced in Section 7.2. We find that the dynamic heuristic not only achieves the optimal performance (in the long-run average cost), but also requires minimal computational efforts, especially when comparing to the one-step improvement heuristic which relies on Monte-Carlo simulation of future sample paths. The heuristic policies we compare in the simulation include:

1. **Dynamic policy** from the algorithm developed in Section 6.
2. **Static policy** from Section 5.2.
3. **Myopic policy**: At each decision epoch, we rank patients using the strong dominance criterion, but we ignore the future cost. This policy can be seen as a special case of the dynamic policy but with the tuning parameters $\tilde{\alpha} = \tilde{\beta} = 0$ in the cost-to-go in (26).
4. **One-step improvement** based on static policy: we obtain an approximation of the cost-to-go function through Monte-Carlo simulation, where on each future day of each simulated sample path, we implement the (optimal) static threshold policy. Then we obtain the number of patients today by minimizing the one-period risk and the approximated cost-to-go by averaging over all simulated sample paths.

Table 3 compares the average costs of these four policies when $C = 1$, $R = 3$, $\Lambda = 6.52$, and $N = 33, 40, 54$. The number in the parenthesis shows the gap from the dynamic policy for the corresponding policy. We can see that the dynamic policy achieves a better performance comparing to the other three. In addition, the one-step improvement policy requires a much longer computation time: it normally requires 10-15 minutes to finish one replication of $10^4$ simulated days, whereas the dynamic policy only requires 0.2-0.3 seconds. All experiments are run on the same computational platform: MacBook Pro with 2 GHz Intel Core i5 processor and 8 GB memory.

Table 4 compares the average costs of these four policies in a setting without readmission arrivals. We set $\Lambda = 8$ and still use $C = 1$, $R = 3$, and $N = 33, 40, 54$. When there is no readmission arrival, we can see that the dynamic policy and the one-step improvement policy achieve similar performance,
both producing the lowest cost. The dynamic policy still gains significant advantages in the computational time comparing to the one-step improvement policy (0.2-0.3 second versus 10-15 minutes for $10^4$ simulated days). The performance of the myopic policy is much worse in this setting.

<table>
<thead>
<tr>
<th>N=54</th>
<th>Dynamic</th>
<th>Static</th>
<th>Myopic</th>
<th>One-step</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.64</td>
<td>3.91</td>
<td>3.78</td>
<td>3.94</td>
</tr>
<tr>
<td></td>
<td>(7.6%)</td>
<td>(3.9%)</td>
<td>(8.3%)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N=40</th>
<th>Dynamic</th>
<th>Static</th>
<th>Myopic</th>
<th>One-step</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4.56</td>
<td>4.75</td>
<td>4.83</td>
<td>4.88</td>
</tr>
<tr>
<td></td>
<td>(4.3%)</td>
<td>(5.9%)</td>
<td>(7.0%)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N=33</th>
<th>Dynamic</th>
<th>Static</th>
<th>Myopic</th>
<th>One-step</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5.27</td>
<td>5.36</td>
<td>5.72</td>
<td>5.60</td>
</tr>
<tr>
<td></td>
<td>(1.8%)</td>
<td>(8.6%)</td>
<td>(6.4%)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3 Average cost from simulation: comparing heuristic policies. The experiment settings are the same as the baseline introduced in Section 7.2, with $C = 1$, $R = 3$, $\Lambda = 6.52$, and $N = 33, 40, 54$.

<table>
<thead>
<tr>
<th>N=54</th>
<th>Dynamic</th>
<th>Static</th>
<th>Myopic</th>
<th>One-step</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.70</td>
<td>3.90</td>
<td>4.57</td>
<td>3.70</td>
</tr>
<tr>
<td></td>
<td>(5.6%)</td>
<td>(23.6%)</td>
<td>(-0.1%)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N=40</th>
<th>Dynamic</th>
<th>Static</th>
<th>Myopic</th>
<th>One-step</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4.31</td>
<td>4.49</td>
<td>5.27</td>
<td>4.32</td>
</tr>
<tr>
<td></td>
<td>(4.3%)</td>
<td>(22.4%)</td>
<td>(0.3%)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N=33</th>
<th>Dynamic</th>
<th>Static</th>
<th>Myopic</th>
<th>One-step</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4.72</td>
<td>4.89</td>
<td>5.74</td>
<td>4.73</td>
</tr>
<tr>
<td></td>
<td>(3.6%)</td>
<td>(21.6%)</td>
<td>(0.1%)</td>
<td></td>
</tr>
</tbody>
</table>

Table 4 Average cost from simulation: no readmission arrivals. The experiment settings are the same as that in Table 3, except that $\Lambda = 8$.

D.3. Readmission arrival distributions

Figure 15 shows the distributions of daily readmission arrivals from simulation experiments. Two policies are used: the dynamic heuristic policy and the static threshold policy. The experiment settings are the same as the baseline introduced in Section 7.2. We can see that the readmission arrivals have almost the same distributions under static and dynamic policies.

![Figure 15](image)
D.4. Daily occupancy under different discharge policies

Figure 16 shows the daily occupancy on different days of a week from simulation experiments. The experiment settings are the same as the baseline introduced in Section 7.2. In subplots (a) and (b), we use the empirical day-of-week arrival pattern estimated from our dataset, and set $N = 40$ and 33, respectively, to test different system loads. In subplot (c), we set $N = 40$ but use a hypothetical arrival pattern that has a larger variability than the empirical one. Three discharge policies are compared in each subplot: the dynamic heuristic policy, the static threshold policy, and the empirical policy that mimics the historical discharge behavior. We can see that the dynamic policy helps smooth the daily occupancy similarly to the results of Section 7.1, whereas the static or empirical policy exhibits little to no smoothing behavior. Though the thresholds are optimized in the static policy, this result suggests that not all optimized discharge policies result in occupancy smoothing. Hence, the dynamic policy is increasingly valuable along multiple dimensions as the variability in the arrival process increases, providing both better performance and greater occupancy smoothing.

Appendix E: Model extensions

E.1. Differentiating patient stages

We consider an extension of the basic model, where each patient goes through two stages capturing the recovering status during her stay: a critical stage, and a stable stage. Similar to Berk and Moinzadeh (1998), we assume that a patient can only be discharged when she is in the stable stage. See Figure 17 for an illustration. We do not differentiate the class when patients are in the critical stage since we are not going to discharging them, but we differentiate the patient classes and LOS when they enter the stable stage. We further assume that the readmission risk is not affected by the time a patient spent in the critical stage but only depends on the time she is in the stable stage.

For this extended model, we modify the state, action, and transition dynamics as follows. Let $Y_k^c$ denote the total number of patients in the critical stage on day $k$, and let $X_k^j = (X_k^{1,j}, \ldots, X_k^{M,j})$ denote the number of patients from each class who have spent $j$ days in the stable stage. The state on each day $k$ is captured by the vector $(Y_k^c, X_k^0, \ldots, X_k^j)$. Consequently, the action is captured by the vector $(D_k^0, \ldots, D_k^j)$ where we do not discharge any of the $Y_k^c$ patients.
The state transition dynamics is modified as follows:

\begin{align}
Y_{k+1}^C &= Y_k^C - \sum_m T_k^m + A_k \tag{53} \\
X_{k+1}^{m,0} &= T_k^m, \tag{54} \\
X_{k+1}^{m,j} &= X_{k}^{m,j-1} - D_{k}^{m,j-1}, \quad j = 1, \ldots, J, \tag{55}
\end{align}

where \(A_k\) denotes the number of arrivals including readmissions, and \((T_k^1, \ldots, T_k^M) \sim \text{multinom}(Y_k^C, \vec{\mu})\) denotes the number of critical patients who become a class \(m\), stable patients in day \(k\), which follows multinomial distribution with probability \(\vec{\mu} = (\mu_1, \ldots, \mu_M)\).

It is straightforward to show that the structural properties from Section 5.1 extend to this case by considering only patients that are in the stable stage, treating patients in the critical stage only as exogenous contributions to the queue length. The interchange argument follows directly since we can only consider interchanging two patients that are dischargeable at any given time period.

### E.2. Randomized static policy

We consider a modification of the static discharge policy. For a given set of thresholds \((l_1, \ldots, l_M)\), we still discharge class \(m\) patients whose LOS has reached \(l_m\) days. However, for class \(m\) patients who spent \(j < l_m\) days, we now assume that they also have a probability \(d_{m,j}\) to be discharged, \(j = 0, \ldots, l_m - 1\). We focus on the setting where there is no class change, i.e., \(\delta_{m,m} = 1\) in (9).

Similar to the basic static-threshold setting, we need to calculate \(\mathbb{E}[D^m]\) (the daily number of discharges) and \(\mathbb{E}[Q]\) in the steady state, and then solve an optimization problem similar to (19). To get both \(\mathbb{E}[D^m]\) and \(\mathbb{E}[Q]\), we need to first calculate \(X_{k}^{m,j}\) for \(j = 0, \ldots, l_m - 1\). Note that a patient with LOS \(j\) days means that she has not been discharged since arrival \(j\) days ago, and this “survival” probability equals \((1 - d_{m,0})(1 - d_{m,1}) \cdots (1 - d_{m,j-1})\). Leveraging the thinning property of Poisson random variable, we know that \(X_{k}^{m,j}\) follows a Poisson distribution with mean \(\mathbb{E}[D^m] \prod_{t=0}^{j-1}(1 - d_{m,t})\), since \(\mathbb{E}[D^m]\) is the daily throughput of class \(m\) patients when the system is in the steady state.

Then, to get \(\mathbb{E}[D^m]\) we use the following set of flow-balance equations similar to (20). The expected arrivals from readmission patients equals

\[ \sum_{j=0}^{l_m} r(m,j) \cdot d_{m,j} \mathbb{E}[X_{k}^{m,j}] = \sum_{j=0}^{l_m} r(m,j) \cdot d_{m,j} \mathbb{E}[D^m] \prod_{t=0}^{j-1}(1 - d_{m,t}). \]

Consequently, the flow-balance equation becomes

\[ \mathbb{E}[D^m] = \Lambda_m + \mathbb{E}[D^m] \sum_{j=0}^{l_m} r(m,j) d_{m,j} \prod_{t=0}^{j-1}(1 - d_{m,t}). \]

The remaining procedure to get the optimal thresholds \(l_m\’s\) as well as the probability \(d_{m,j}\’s\) is similar to that in the basic static-threshold setting.
Appendix F: Enlarged figures for the implemented tool
Figure 20  Portal for changing default parameter setting.

Figure 21  Screen shot for the predicted readmission timing for post-discharge monitoring.