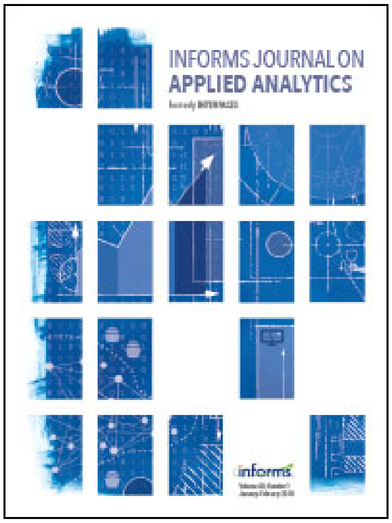


INFORMS Journal on Applied Analytics

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Optimization Automates Emergency Department Nurse Scheduling at Hartford Hospital

Liangyuan Na, Jean Pauphilet, Ali Haddad-Sisakht, Louis Raison, Audrey Silver, Patricia Veronneau, Nicole Vogt, Dimitris Bertsimas

informs
https://pubsonline.informs.org/journal/ijaa

INFORMS JOURNAL ON APPLIED ANALYTICS
Vol. 54, No. 6, November–December 2024, pp. 553–574
ISSN 2644-0865 (print), ISSN 2644-0873 (online)

Optimization Automates Emergency Department Nurse Scheduling at Hartford Hospital

Liangyuan Na,[a,*] Jean Pauphilet,[b] Ali Haddad-Sisakht,[c] Louis Raison,[c] Audrey Silver,[d] Patricia Veronneau,[d] Nicole Vogt,[d] Dimitris Bertsimas[e]

[a] Operations Research Center, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139; [b] Management Science and Operations, London Business School, London NW1 4SA, United Kingdom; [c] Dynamic Ideas LLC, Waltham, Massachusetts 02452; [d] Hartford HealthCare, Hartford, Connecticut 06103; [e] Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

*Corresponding author

Contact: nliangyuan@gmail.com, https://orcid.org/0000-0002-2900-9899 (LN); jpauphilet@london.edu, https://orcid.org/0000-0001-6352-0984 (JP); ali@dynideas.com (AH-S); lraison@dynideas.com (LR); audrey.silver@hhchealth.org (AS); patricia.veronneau@hhchealth.org (PV); nicole.vogt@hhchealth.org (NV); dbertsim@mit.edu, https://orcid.org/0000-0002-1985-1003 (DB)

Received: September 24, 2023
Revised: January 12, 2024
Accepted: April 9, 2024
Published Online in Articles in Advance: July 10, 2024

https://doi.org/10.1287/inte.2023.0071

Abstract. To optimize nurse staffing in the emergency department (ED), Hartford Hospital has been collaborating with academics and consultants to schedule nurse shifts over each six-week staffing cycle. We develop and implement two-phase optimization models: a robust optimization model to find optimal staffing levels given the uncertainty in patient demands, followed by a pair of mixed-integer problems to generate individual schedules including work, trainee, and preceptor shifts for each nurse. Our approach leads to less costly (5%–8%) staffing with better coverage of patient care (8%–25%) and higher nurse satisfaction (5%). Moreover, nurses can work fewer shifts on weekends (17%), holidays (14%), and overtime (85%), as well as be assigned to more diverse positions (3.6) and more daily training opportunities (0.95). We implement our framework into an automated end-to-end scheduling optimization software, deployed for use at Hartford Hospital since March 2023. The software collects preferences from more than 200 ED nurses and enables managers to optimize schedules with guided dynamic adjustments. This transformative implementation streamlines a previously labor-expensive staffing process (currently taking more than 88 manual hours per cycle) and delivers schedules that are more suitable for patients and nurses together, with an annual projected cost saving of around $720,000.

History: This paper was refereed.

Keywords: robust optimization • mixed integer optimization • software automation • nurse scheduling • emergency department

Introduction

Emergency departments (EDs) are a crucial component of hospitals, providing urgent care to patients in need. However, ED overcrowding can lead to increased waiting times, compromise care, and result in adverse outcomes for patients, such as dissatisfaction and increased mortality rates (Bernstein et al. 2009). Additionally, ED congestion can affect patient flows into the hospital, further impacting care delivery (Elder et al. 2015).

Nurse staffing is vital for patients, especially ED patients in critical condition, having a direct impact on their health outcomes. Among other evidence, more hours of care provided by registered nurses (RNs), a lower nurse workload, and a higher nurse skill set have been associated with reduced mortality rate and shorter hospital stays (Needleman et al. 2002, Aiken et al. 2014, Twigg et al. 2019). For the ED as a whole, a shortage of nurses also negatively impacts throughput metrics, such as length of stay or abandonment (i.e., patient leaving without being seen; Ramsey et al. 2018). However, healthcare systems are facing increasing difficulties to meet appropriate nurse-to-patient ratios, owing to nurse shortages, burnout, and dissatisfaction (Aiken et al. 2002), as exacerbated since the COVID-19 pandemic (Peters 2023). For example, systemic overtime and dissatisfaction with schedule flexibility are common reasons for nurses to leave the profession (Leineweber et al. 2016).

One challenge is to schedule the "right" number of nurses—enough to accommodate future patient demand without overstaffing and wasting precious and limited nursing resources. Another challenge is to increase the satisfaction of each nurse regarding individual schedule by taking into account individual preferences and requirements. To improve both patients' and nurses' experiences, we use optimization for nurse staffing and scheduling in the ED and address the aforementioned challenges. Our final objective is to positively impact nurses and patients as well as the overall performance of the ED and the hospital, which we achieve through a tight and long-standing collaboration between ED practitioners and academics.

ED Nurse Staffing Problem at Hartford Hospital

Hartford HealthCare is Connecticut's largest healthcare network, operating in more than 400 locations and providing a broad range of services. Hartford HealthCare comprises seven hospitals, with more than 2,400 inpatient beds. Hartford HealthCare's flagship facility is Hartford Hospital (HH), a teaching hospital and tertiary care center with 867 beds, in partnership with the University of Connecticut School of Medicine. With more than 160 years of experience, HH is renowned for its exceptional performance in various procedures, conditions, and specialties. The hospital serves more than 40,000 discharges and 100,000 ED visits annually.

Nurses working in the ED are typically assigned to one of 13 "positions." Five positions cover managerial and logistical duties: clinical leader (CNL), first nurse, resource nurse, triage nurse, and front end provider (FEP). The remaining eight positions, also referred to as "pods," correspond to treating different categories of patients. There are four main pods (blue, green, and orange pods and their hallways) in which most of the patients are treated: one red pod for resuscitation/emergent patients, one purple pod for behavioral patients, an iTrack pod for less urgent patients, and one ED observation unit (EDOU) for observation patients. This organization was disrupted during the early surge of COVID-19: HH opened up a new position, a "trailer pod," to treat COVID-19 patients specifically and stopped using the hallways to reduce patient cross-infection.

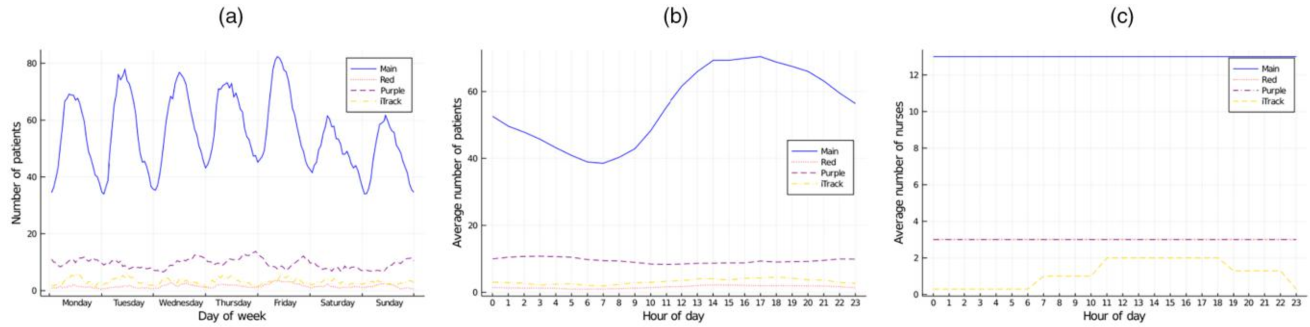
Not every nurse can work at any position. At HH, nurses are classified into nine different tiers based on their experience and qualifications, which dictate the positions they can fill. In addition, based on their years of experience, nurses can receive preferential treatment regarding weekend schedules (see Anderson et al. (2022) for another example of "Armstrong" seniority requirements in nurse scheduling).

Each day is divided into three 12-hour shifts (7 a.m.–7 p.m., 11 a.m.–11 p.m., and 7 p.m.–7 a.m.). Accordingly, ED nurse scheduling, the problem of assigning nurses to pods and shifts, can be divided into two phases. First, one needs to estimate the "right" number of nurses needed at each pod for each future shift (i.e., staffing targets). Second, one needs to find a reasonable assignment of individual nurses to shifts to achieve these targets. Furthermore, at this phase, the ED leadership might want to schedule additional training shifts, during which a trainee seconds a preceptor.

Regarding the first phase, prior to our effort, HH was using predetermined staffing targets. These targets were the same for all days (among others, they did not differentiate between weekdays and weekends). For most positions (i.e., all positions except triage and iTrack), the staffing targets were also constant across shifts. However, patient arrivals—and hence, the need for ED nurses—vary across days and within each day. Figure 1 demonstrates such inconsistency between fluctuating demand levels and fixed staffing levels (see more in the appendix section Demand Patterns and Standard Scheduling). This contrast can lead to overstaffing at some times, increasing hospital staffing costs, and understaffing at other times, compromising patient care quality. COVID-19 has further complicated the situation by disrupting patient demand patterns and increasing employee attrition.

Regarding the second phase, ED nurse schedules were planned for six-week periods according to the following timeline: Before each staffing cycle, nurses enter their preferences and availability. Schedulers and nurse managers manually generate a schedule that aims to balance staffing levels and preferences between nurses. In addition, the ED leadership manually adds training sessions to the schedule whenever possible. After the schedule is announced, nurses can ask for amendments to better satisfy individual preferences and increase fairness among nurses. In turn, managers and schedulers need to constantly recompute schedules to accommodate as many requests as possible. To announce the

Figure 1. (Color online) Patterns of Average Patient Demand vs. Standard Staffing Levels



Notes. (a) Day-of-week demand. (b) Hour-of-day demand. (c) Daily schedule.

schedule, scheduling managers take an additional step to manually convert the schedule of approximately 3,600 nurse shifts (on average, three weekly shifts per nurse for 200 nurses over six weeks) into a "team sheet" of a specified format. HH estimates that the entire scheduling process takes managers and schedulers 88 hours of manual work per cycle and can be prone to errors.

Since 2020, we, a collaboration of nursing providers, academics, and data consultants, have been working to improve the ED nurse scheduling process at HH. We develop an integrated approach consisting of a two-phase optimization methodology and a software implementation. First, we use a data-driven optimization approach that allocates limited staffing resources cost-effectively and determines sufficient staffing targets to reach appropriate nurse-to-patient ratios, taking into account variability in patient arrivals and needs. Second, we generate an optimized nurse-to-shift allocation that aims to match nurses' individual requests and preferences, increase nurse satisfaction, and provide opportunities to schedule training shifts. Finally, we develop scheduling software that relieves manual labors and operations burdens for nurse managers and schedulers. Our tool was partially used at HH ED in March 2023 (in preparation for the staffing cycle of April 9 to May 20) and started to be fully used for future cycles. Overall, our decision-support tool leverages a combination of data, optimization, and software to achieve a holistic improvement for ED staffing.

Related Work

The operations research literature contains numerous methods for ED operations, which we summarize here. We refer to Saghafian et al. (2015) for a comprehensive review.

Simulation-based approaches use simulations of the ED environment to evaluate alternative scenarios and, combined with optimization, to better allocate resources (Chen and Wang 2016), staff (e.g., doctors, laboratory technicians, and nurses; Ahmed and Alkhamis 2009), and particularly nurses (Draeger 1992). Alternatively, patient arrivals can be well approximated by Poisson processes with nonstationary (Kim and Whitt 2014) or uncertain (Maman 2009) arrival rates. Hence, queuing approaches have been used to model the ED and address staffing problems such as, recently, dynamic shift assignment (Chan et al. 2021) or surge staffing (Hu et al. 2024). There is a rich literature on formulating personnel scheduling problems in general (Brucker et al. 2011)—and nurse scheduling problems in particular (Svirsko et al. 2019)—as formal optimization problems with constraints and often multiple objectives (e.g., workload, staffing costs, individual preferences; Ang et al. 2018, Mohammadian et al. 2019, Hamid et al. 2020, Rerkjirattikal et al. 2022), solved by iteratively solving for each objective or using goal programming techniques.

Optimization methods can account for uncertainty in their input parameters, here in patient arrivals/needs and in nurse availability. Lim and Mobasher (2011) and Van Hulst et al. (2017), for example, use robust optimization to generate workforce plannings that are robust against uncertain per-patient workload. In addition to demand uncertainty, frequent change in nurse availability due to sickness or absenteeism is another bottleneck for the implementation of optimized schedules in practice. Clark et al. (2015) highlight the importance and challenges of shift rescheduling and advocates for mathematical models to support the rescheduling process. For example, Wickert et al. (2019) propose an integer optimization formulation for the nurse shift rescheduling problem. In this work, we directly account for variability in patient arrivals and needs in the first phase of our optimization approach, by proposing a robust optimization problem to find appropriate staffing targets. Although we do not directly account for uncertainty in nurse availability, we aim to maximize nurse engagement and adherence to the schedule by developing the second phase of our approach, which takes into account nurses' individual preferences. In addition, the software interface we develop allows nurses to update their availability and preferences directly in the system and enables ED managers to dynamically recompute schedules with the latest input from nurses, hence materializing the rescheduling process advocated by Clark et al. (2015).

Most related to our work is Ang et al. (2018), who develop a mixed-integer sequential goal programming model to optimize nurse-to-patient ratios and shift preferences, among others. They integrate their model into an online decision-support system to ease practitioners' adoption. Their paper provoked a vibrant discussion in the nursing community, with critics calling such solutions "unimplementable" (Park et al. 2022). Despite the abundance of academic work in the past decade on nurse scheduling, there is still a gap between research and practice, with supposedly only 30% of nurse scheduling models from research ever being implemented, let alone still being used today (Kellogg and Walczak 2007).

Another obstacle to implementation is the lack of mathematical modeling and optimization skills among nurses (Park et al. 2022). To bridge this gap and advance nursing science, we collaborate closely with nurses and nurse managers to collectively reach an optimization formulation that can be realistically implemented. We also develop a user-friendly end-to-end software interface for nurses to use, and we train nurse leadership to run the scheduling optimization on their own. Our model-software integration, jointly with trust built with the medical team, results in the successful deployment at HH.

Main Contributions

We summarize our contributions as follows.

1. We develop a two-phase optimization approach for ED nurse staffing: In the first phase, we leverage robust optimization and historical data to optimize aggregate nurse staffing levels. Then, taking these aggregate optimized staffing levels as inputs, we develop mixed-integer optimization models to generate an individual-level schedule that prioritizes individual nurse preferences and training opportunities.

2. On computational experiments, we demonstrate the benefit of our two-phase optimization approach: our first-phase optimization model reduces costs by 5%–8% during overstaffing periods and reduces insufficient nurse-to-patient-ratio coverage by 8%–25% during understaffing periods. In addition, we conduct various experiments that compare different modeling variants, illustrate the trade-offs between the schedule's different objectives, and provide strategic insights for ED leadership. The second-stage model for individual-level schedules improves individual nurse satisfaction by 5% while increasing position diversity, training opportunities, and fairness.

3. We implement our optimization models into an end-to-end scheduling software that covers all staffing steps, from configuration preparation and input collection to schedule generation and output production. For every scheduling cycle, nearly 200 ED nurses enter their availability and preferences into the system. Through an interactive interface, nurse managers can run the optimization models on their own to generate, edit, and announce schedules in a few clicks. In addition to the benefit of optimization, this unifying interface and software has accelerated our collaboration, revolutionized ED nurse scheduling, and reduced manual burdens to a minimum.

ED Aggregate Staffing Optimization

In the first phase, we develop a robust optimization model to determine the right number of nurses to staff in the ED, at each position and each shift, for the next six weeks.

Optimization Model Overview

Our key decision variables are the number of nurses to staff for each position j, each shift i, and each day d of the next 42-day period. We differentiate nurses based on their tier q and their weekend group g (some nurses can work every other weekend, whereas others work every third weekend only). Therefore, we denote z_{qgjid} the number of nurses of tier q and weekend group g who are staffed at position j on the shift i of day d.

We then evaluate our decision based on three concurrent objectives. First, based on historical data, we can estimate the number of nurses required to treat patients present in the ED, for each position and each hour. We call *insufficiency* the number of nurse shifts missing—that is, the gap between the number of nurse shifts needed and the number staffed. Because the number of nurse shifts needed is uncertain, we adopt a robust perspective and try to minimize the worst-case insufficiency, in which the worst case is taken over a range of possible scenarios we construct from data, as described in the following paragraph. Second, we want to minimize staffing costs, which can be approximated by the sum of scheduled nurse shifts, $\sum_{q,g,j,i,d} z_{qgjid}$ (one could easily add weights based on tiers, for example). Finally, to increase the adoption of our recommendation, we want to minimize the number of changes recommended compared with the current schedule. We incorporate these three objectives into a single objective function with weight parameters, which we tune based on the desirable tradeoffs.

Our model can incorporate many operational constraints. For example, we incorporate constraints to account for the maximum number of weekly working hours for each nurse and their weekend schedule patterns or to impose some week-over-week regularity in staffing levels. We detail our optimization formulation, its objective, and constraints in the section Aggregate Optimization Formulation.

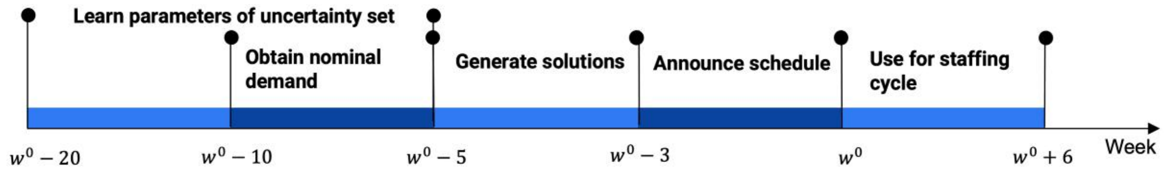
Estimating and Modeling Demand

Using historical data, we estimate how many nurses are needed for each position and each hour.

Data Collection and Processing. HH's information technology system daily transfers records of ED patients to our data repository on Amazon Web Services using a secure file transfer protocol server. The data record detailed information about each patient such as Emergency Severity Index (i.e., acuity level), service, ICD-10 code, and discharge disposition (e.g., admitted to intensive care unit, surgery, interventional radiology, or discharge). Based on current ED practice, we use this information to match every patient to an appropriate pod and a target nurse-to-patient ratio (1:1 or 1:2 for those in the red pod, 1:7 in the purple pod, 1:12 in iTrack, and 1:5 in the remaining pods). Hence, at any point in time (in practice, we aggregate the data at an hourly level), we compute the number of patients present in each pod and the number of nurses needed.

Modeling Uncertainty. Let us denote $\tilde{h}_{jws}$ the number of nurses needed at position j on week $w \in \{1, \dots, 6\}$ at hour $s \in \{1, \dots, 24 \times 7 = 168\}$. The schedule for the next six-week cycle is announced three weeks in advance. Hence, we start formulating, calibrating, and solving our optimization problem five weeks before the beginning of the cycle, as described in Figure 2. At this time, we estimate the number of nurses needed in the past

Figure 2. (Color online) Timeline of Uncertain Demand Modeling for Every Staffing Cycle



six weeks and use it as our baseline estimate for demand in the coming six weeks, $\overline{h}_{jws}$. The actual demand for nurses $\tilde{h}_{jws}$, however, might deviate from this baseline estimate, in part because it uses data that will be 10 weeks old by the time the schedule is implemented. Accordingly, we consider all demand vectors contained in an uncertainty set $\mathcal{U}$ and calibrate the size of the uncertainty set based on the distance between demand vectors that are 10 weeks apart, using data from w_0-20. The full uncertainty set construction is described in the section Aggregate Model Data.

Executing the Model

Solving the Robust Optimization Model. Because both the decision variables and uncertainties are integers, duality theory is not applicable, and thus, a closed-form robust counterpart cannot be obtained. Hence, we use a cutting-plane approach to identify a subset of the most restrictive constraints and approximate the worst-case objective value among the uncertainty set. All models in this work are implemented using the JuMP package (Dunning et al. 2017) in Julia programming language (Bezanson et al. 2017) and are solved by the Gurobi solver (Gurobi Optimization, LLC 2023). Details of solving the model, such as constraint linearization, algorithmic and programming implementation, and specifications, are included in the section Solving the Robust Optimization Model.

Parameter Tuning. In preparation for each upcoming staffing cycle starting on week w^0, we can manually tune a combination of parameters, including two weights in the objective terms, the size of the uncertainty set, and relative penalty for each week's insufficiency (penalize more on later weeks' demand from the uncertainty set). To validate our choices, we examine the performance of different combinations of parameters over a validation period of w^0-10 to w^0-5. By comparing the values of the objective terms, we determine the optimal combination of parameters. For comparative purposes, we also consider a deterministic model in which we only use the nominal demand instead of incorporating the uncertainty set. We demonstrate some of the process in the Results section.

Quantifying Strategic Decisions

Our model not only optimizes staffing levels for each cycle but can also help nurse leadership inform some strategic decisions for the ED. For example, we can evaluate the impact of having staffing targets that are the same for all days or the same from one week to another. We can simulate the benefit of creating a fourth shift (e.g., 2 p.m.–2 a.m.). Details on incorporating these variations into the model are presented in the section Model Variants for Informing Strategic Decisions.

ED Individual Scheduling Optimization

For the second phase, we develop another optimization model to schedule individual nurses based on the recommended aggregate staffing levels.

Optimization Model Overview

Our core decision variables are individual nurse assignments—that is, when and where each nurse works (and trains). We use binary variables $b_{\ell jid}$, $r_{\ell jid}$, and $p_{\ell jid}$ to denote whether nurse ℓ works, trains, or mentors, respectively, at position j during shift i on day d.

Our primary objective is to maximize nurse satisfaction across various metrics, such as individual preference on dates, times, and shift patterns, diversity in pod assignments, and reduction in weekend, holiday, and overtime shifts. Meanwhile, we penalize shortages and surpluses in aggregate staffing levels while rewarding scheduled training shifts. We jointly optimize for a combination of four objective terms with adjustable tradeoffs between the components.

In addition, we include a wide range of constraints. Some are to keep track of the different objective values, such as nurse satisfaction scores and deviations from aggregate staffing levels. Other constraints enable a feasible assignment, subjective to labor regulations, ED requirements, and eligibility, among others. We present a detailed formulation of the optimization model in the section Individual Shift Scheduling Model.

Individual Model Data

One main input of the second phase's model is the output from the first phase, namely the number of nurses of each tier/group to schedule for each position during each shift on each day. In addition, the model takes into account a range of individual preferences and the availability of each nurse:

- Feasibility and preference to work at each shift type
- Total number of weekly shifts, paid time off (PTO), and education time of each week
- Weekend group
- Preference and availability to work on each day
- Preference for different work patterns. Typically, nurses can express a preference for one of three patterns:
    1. All shifts in a row (e.g., working on Monday, Tuesday, and Wednesday consecutively)
    2. Every other day (e.g., Monday on, Tuesday off, Wednesday on, …)
    3. Two on–two off (e.g., Monday and Tuesday on, Wednesday and Thursday off, …)
- Eligibility to be a trainee or a preceptor (training the trainees) at each position

Integer Optimization Model on Preceptor Scheduling

After obtaining the optimal solution for scheduled work and trainee shifts, we develop and solve another mixed-integer linear optimization model to schedule preceptor shifts to train the corresponding trainees. The overall goal is to distribute trainees more evenly among eligible preceptors. Thus, we set the objective function to minimize the maximum number of preceptor shifts assigned to each nurse. We formulate the preceptor scheduling optimization model in the section Individual Preceptor Scheduling Model.

Flexible Options for Nurse Leadership

We provide various options for ED nurse leadership to generate alternative schedules based on their preferences. These options include

- The option to forbid a change of shift type from each nurse's preassigned shift type
- The ability to exclude trainee scheduling
- The option to decompose the model into ED and EDOU and solve for each of the two partitions separately
- Control over the priority importance of different terms in the objective function by adjusting the weighting factors. Currently, the order of importance is shortage, satisfaction, unassigned/surplus, and then training. If training is not intended to be scheduled, it can be excluded by setting its weight to zero.
- The option to penalize shortages differently for each position j through weights w_j^{shortage}. Currently, positions that require more senior nurse tiers (such as CNL and triage) have higher penalization weight for the shortage.
- Fine-tuning of the relative importance of satisfaction score metrics using λ's. To reduce model solving time, the preferred pattern and/or unpreferred pattern term(s) can be turned off by assigning a weight of zero if desired.
- Inclusion of fairness among nurses as a hard constraint controlling worst-case penalty score, or as a soft constraint by adding the worst-case score term in the objective function with a corresponding weight factor. Alternatively, fairness can be excluded altogether, which could also reduce solving time.

Results

In this section, we illustrate the benefits of our two-phase optimization approach. We decide to report these results chronologically, following the timeline and progress of the collaboration, for three reasons. First, we believe other researchers might find it insightful to know the successive iterations and milestones our collaboration has gone through. Second, because of the COVID-19 pandemic, the number of ED visits significantly changed between the first (end of 2020) and the latest (end of 2022) evaluations of our model. Accordingly, the value of optimization switched from reducing staffing costs to reducing insufficiency risk. Finally, at this stage in the project, we had not developed our user-friendly interface yet. Hence, iterations on the models between the research and nursing teams were slow and primarily paced by the six-week staffing cycles. These frictions in our collaboration were the prime motivation to develop and deploy software for nurses to use directly.

End of 2020: Reducing Staffing Costs During Low-Demand Periods

We first evaluated the benefit of our first-stage optimization problem (staffing levels) at the end of 2020, using data from October 26 to December 20, 2020. In particular, we generated three instances of six-week staffing cycles by considering three overlapping six-week periods (beginning one week apart) from October 26–December 6 to November 9–December 20.

We measure the quality of our staffing levels based on two metrics:

- *Cost*: the total number of nurses staffed, on average per day. This cost is not subject to any uncertainty.
- *Insufficiency*: the number of extra nurse shifts needed to satisfy the demand, on average per day. We obtain this number by computing the number of extra nurses needed to satisfy the target nurse-to-patient ratios on the realized demand at each hour of the day and sum across all hours in a shift and all shifts in a day. We then average over all days in the six-week cycle. For example, an average insufficiency of 0.17, as is the case for the historical schedule (Table 1), means

Table 1. Results of Average Outcomes (±Standard Deviation) from Six Approaches

Schedule approach	Shift types	Daily cost	Cost reduction (%)	Daily insufficiency	Insufficiency reduction (%)
Current	7–7, 11a–11p	56	—	0.17 (±0.06)	—
Oracle	7–7, 11a–11p	49.81 (±0.46)	11.05 (±0.82)	0.02 (±0.00)	86.96 (±0.33)
Nonrobust	7–7, 11a–11p	50.42 (±0.38)	9.95 (±0.67)	0.21 (±0.04)	−22.55 (±79.57)
Robust (a)	7–7, 11a–11p	51.33 (±0.50)	8.33 (±0.90)	0.12 (±0.03)	27.70 (±81.31)
Robust (b)	7–7, 12p–12a	50.57 (±0.49)	9.69 (±0.88)	0.15 (±0.07)	13.18 (±72.64)
Robust (c)	7–7, 11a–11p, 2p–2a	50.57 (±0.38)	9.69 (±0.67)	0.15 (±0.07)	11.45 (±74.95)

Note. 7–7, both the 7 a.m.–7 p.m. shift and the 7 p.m.–7 a.m. shift.

that we need to add on average 0.17 nurse shifts per day to satisfy the target nurse-to-patient ratios.

We aim to reduce both costs and insufficiency.

Table 1 reports the daily staffing cost and insufficiency for different policies, as well as their relative reductions compared with the current staffing levels. The current staffing levels require 56 nurse shifts per day and lead to an average insufficiency of 0.17 nurse shifts. As a comparison, we compute the optimal staffing level obtained by an oracle that knows the realized demand for the coming six-week cycle. It reduces daily costs to 49.81 (11.05% reduction) and insufficiency to 0.02 (86.96% reduction), suggesting significant potential for improvement from using analytics on both quality measures. Although the potential to reduce insufficiency is significant in relative terms, we should emphasize that the current schedule experiences low insufficiency levels in absolute terms, so we are primarily interested in reducing costs at this stage. These results are reported under a certain selection of hyperparameter values, and we present metric trade-offs under different parameters in the section Supplementary Results During Low-Demand Periods.

Our robust optimization approach (Robust (a)) recovers part of the benefit from the oracle by both costs and insufficiency, to 51.33 (8.33% reduction) and 0.12 (27.70% reduction), respectively. Compared with the oracle, this optimization model does not use the realized demand for each six-week cycle but rather estimates it using historical data and constructs an uncertainty set around this estimate. Indeed, we observe that the nonrobust version of this model, which takes demand estimates at their face value, reduces costs by a comparable amount (to 50.42) but significantly increases insufficiency compared with the current schedule because it fails to account for variability in nurse needs.

Finally, we leverage our optimization model to inform strategic decisions for the ED nurse leadership. For example, we consider moving the 11 a.m.–11 p.m. shift one hour later (Robust (b)) or introducing a fourth shift, 2 p.m.–2 a.m. (Robust (c)). Compared with Robust (a), we observe that these changes can further reduce costs, with comparable (yet slightly worse) insufficiency. After reviewing these changes with the nurse leadership, we decided to keep the existing shift structure and recommended the robust schedule (a).

We provide details in the section Supplementary Results During Low-Demand Periods that illustrate how the recommended schedule patterns better match demand patterns.

2021: Accounting for Frictions to Implementation

Given the positive results obtained at the end of 2020, our optimization approach gained support from the hospital executives. Hence, in 2021, we worked closely with the nurse leadership to refine the model and anticipate any friction to implementation. In particular, it was during this phase that we started penalizing the number of changes compared with the current schedule in our model for the staffing levels and that we started building the individual nurse-to-shift assignment model.

Because of the flexibility provided by our mathematical model, we were able to incorporate many requirements from the nurse leadership to better align the solution of our model with the needs of the nursing staff. For instance, we fixed the staffing level in the red pod (the one for the most critical patients) to the current level. We introduced the possibility for nurses to "float" (or reassign) between pods—that is, between main pods or from the red pod to the main pods when red pod demand is low. These improvements led to the first staffing level recommendation, displayed in Table 2, Panel A, for the cycle of May 30–July 10, 2021. The optimization-based solution schedules more 11 a.m.–11 p.m. shifts but fewer 7 a.m.–7 p.m., 7 p.m.–7 a.m., and weekend shifts, resulting in an 8% reduction in the total number of shifts and a better match between demand and staffing levels.

However, the nurse leadership was reluctant to implement such a significant reduction in staffing levels, citing various concerns. First, the current ED staffing levels had been in place for decades, thus providing stability and visibility to nurses. To address this concern, we added the third component in our objective function that penalizes the distance between the current and the proposed schedule. Second, despite long and extenuating working hours, nurses tend to prioritize quality of care (i.e., insufficiency) over cost

Table 2. Recommended Changes to Overall Staffing Levels

	Blue			Green			Orange			Purple		
Panel A: Iteration 1												
	7a	11a	7p	7a	11a	7p	7a	11a	7p	7a	11a	7p
Sun.	−1	0	−1	−1	0	−1	−2	0	−2	−1	0	−1
Mon.	−1	+1	−1	0	0	0	−1	+1	−1	0	0	0
Tues.	0	0	0	0	0	0	−1	+1	−1	0	0	0
Wed.	0	0	0	0	0	0	0	0	0	0	0	0
Thurs.	−1	+1	−1	0	0	0	−1	+1	−1	0	0	0
Fri.	0	0	0	0	0	0	−1	0	0	0	0	0
Sat.	−1	+1	−1	−1	0	0	−2	+1	−2	−1	0	0
Panel B: Iteration 2												
	Blue			Green			Orange			Purple		
	7a	11a	7p	7a	11a	7p	7a	11a	7p	7a	11a	7p
Sun.	−1	0	−1	−1	0	−1	−2	0	−2	−1	0	−1
Mon.	0	0	0	0	0	0	0	0	0	0	0	0
Tues.	0	0	0	0	0	0	−1	+1	−1	0	0	0
Wed.	0	0	0	0	0	0	0	0	0	0	0	0
Thurs.	0	0	0	0	0	0	0	0	0	0	0	0
Fri.	0	0	0	0	0	0	−1	0	0	0	0	0
Sat.	−1	+1	−1	−1	0	−1	−1	0	−1	−1	0	0

reduction. Based on this observation, we refined our calibration of the tradeoff parameters in the objective and of the size of the uncertainty set to reflect their preferences and worked on a software interface that allows them to explore different solutions for different configurations and choose the one they deem is best. Third, aware of the disruptions and nonstationarities due to the COVID-19 situation, they raised concerns about the relevance of past data to predict demand in the next six-week period. In response to this concern, we added weights to the definition of insufficiency to allow us to put more emphasis on the most recent data, which should be more predictive of future demand.

After taking into account all these changes, we obtained a second proposal, shown in Table 2, Panel B, with less drastic changes, which made the nurse leadership more comfortable with our tool. These iterative adjustments have undeniably helped build a closer relationship and trust between the research and nursing teams.

2021: Building Individual Schedules

The development of optimized staffing levels justifies the need to automate the design of individual schedules as well, because optimization might introduce more changes from one staffing cycle to the next than the current practice (which uses the same staffing levels for all cycles).

Based on staffing levels from Iteration 2, we solve individual schedules for the cycle of May 30–July 10, 2021, on a roster of 110 ED nurses (excluding per diem nurses) and 26 EDOU nurses based on their preferences. Table 3 compares the standard schedule with the one returned by our optimization model (see the section Supplementary Results During Iteration Periods for additional results with alternative model variations). We observe that the optimization model reduces the number of daily nurse shifts by 5%, with a higher reduction during weekends (17%) and holidays (14%). Additionally, the optimization leads to a remarkable 85% reduction in overtime shifts. Such savings in staffing cost and nurse workload is achieved while keeping the insufficiency level below 0.1%. In addition to cost savings, the optimized schedule has mainly additional practical benefits: It creates more opportunities for training new nurses by scheduling nearly one training shift per day (compared with zero in the current schedule). It provides more diversity to nurses by assigning them to 4.6 different positions across the staffing cycle, compared with 1 currently. Finally, it decreases the individual nurses' dissatisfaction penalty score by 5% while also ensuring fairness among their scores. Overall, our optimization significantly improves operational efficiency and nurse satisfaction in the ED.

2022: Reducing Insufficiency During High-Demand Periods

In 2022, the ED patient volumes recovered to (and exceeded) their pre–COVID-19 levels compared with those in our initial evaluation in 2020. Consequently, HH changed the structure of the ED and decided to

Table 3. Comparison of Standard vs. Optimized Schedule

Metric	Standard	Recommended	Reduction (%)
Work shifts per day	48.3	45.7	5
Per weekend day	48.0	40.0	17
Per holiday	48.5	41.5	14
Overtime per day	1.3	0.2	85
Insufficiency per day	<0.1	<0.1	—
Training shifts per day	0	0.95	—
Different positions per nurse	1	4.6	—
Weighted dissatisfaction score	101	96	5

increase the number of staffed nurses. We now illustrate how the same model as the one we developed in 2020 (after some adjustments to account for new positions and tiers) continues to provide substantial benefits, yet primarily by reducing insufficiency instead of costs.

Figure 3(a) shows the insufficiency versus costs trade-off for different solutions obtained from our optimization models with different weight parameters—that is, varying the relative penalty applied to each objective component and varying the size of the uncertainty set. For simplicity, we do not impose a penalty on the changes compared with the current schedule for these simulations. First, let us observe that, because of higher demand, the insufficiency levels are much higher than previously experienced (around 5 nurse shifts per day versus 0.17). Qualitatively, this curve informs the ED leadership on how additional staffing translates into insufficiency reduction. Based on these simulations, we selected a combination of parameter values (yellow star) that was providing a strict improvement over the current schedule, in terms of both costs and insufficiency. We then fine-tuned the other parameters (penalty on deviation from the current schedule in the objective and weight on the most recent data in the definition of insufficiency) as displayed in Figure 3(b). Our final parameter combination reduces insufficiency by 8% (from 5.08 to 4.66 nurse shifts) with 2.38% lower staffing costs (from 66 to 64.43 daily shifts). We provide details of the parameter tuning process in the section Supplementary Results During High-Demand Periods.

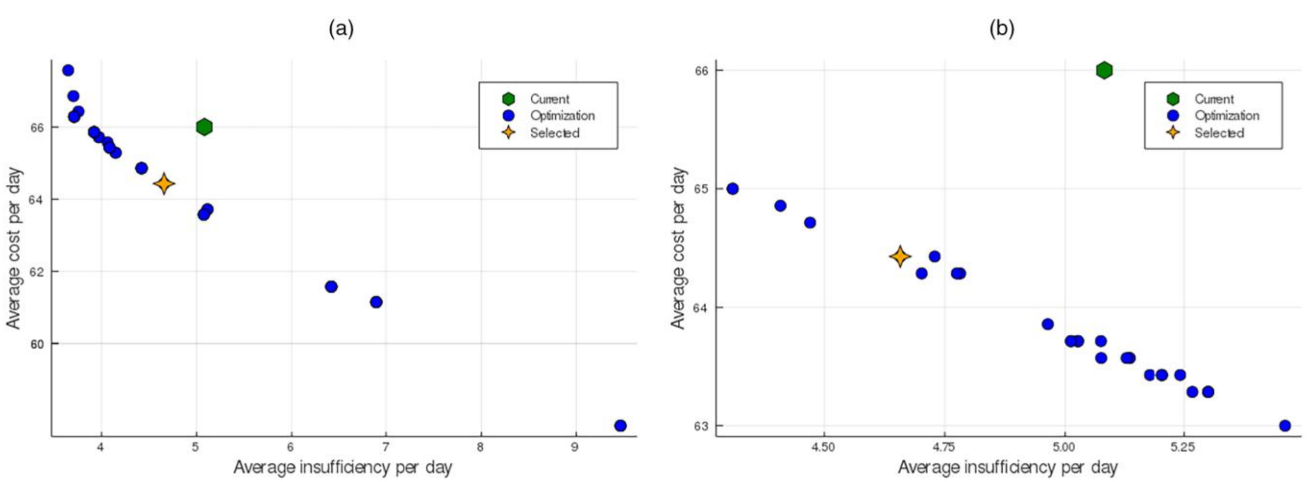
To further facilitate automating the scheduling in adjustment to demand variation, we investigate the benefit of resolving the aggregate model and adjusting the schedule every week during the six-week staffing cycle in the section Supplementary Results During High-Demand Periods.

Implementation

In collaboration with a data consultancy company and a development team, we integrate our models into end-to-end software interface.

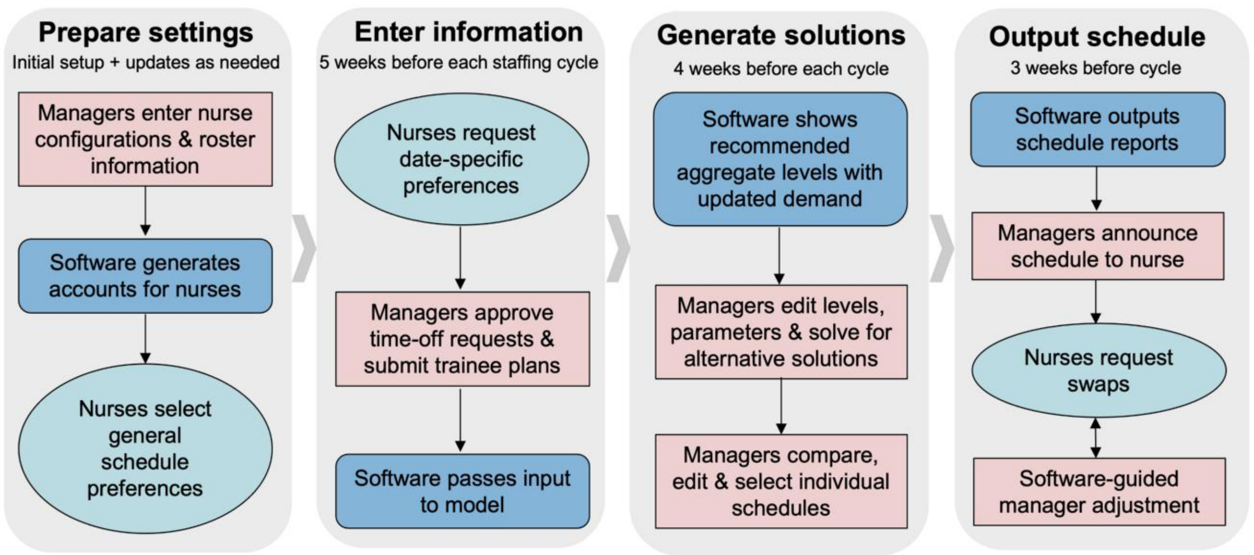
The process, described in Figure 4, outlines the four phases of the scheduling process: preparation, input collection, solution generation, and schedule production.

Figure 3. (Color online) Objective Tradeoffs During Staffing Cycle December 4, 2022, Through January 14, 2023



Notes. (a) Tuning insufficiency parameters (weights and uncertainty set size). (b) Fine-tuning other parameters (penalization on changes and recency).

Figure 4. (Color online) Process of Integrating Decision-Support Software into ED Nurse Staffing



The flowchart depicts the interactions between the software (blue rounded rectangle), ED nurse managers (red rectangle), and individual nurses (green oval) at HH. Our software revolutionizes the ED nurse scheduling process as follows: The process starts with nurse managers setting up general configurations and nurse information in the software, which then enables individual nurses to log into their accounts and enter general preferences. This phase is completed prior to the first use and can be updated whenever necessary, whereas the remaining three phases are performed for every six-week staffing cycle. Approximately five weeks before the start of each cycle, the software collects nurses' and managers' requests specific to the upcoming cycle and then passes them as inputs to the optimization model. Four weeks before the cycle, the software automatically solves the first-stage aggregate model, allowing managers to edit the output aggregate levels and run the program to generate schedules with parameters of their choice. After managers edit, compare, and select the schedules, the software outputs reports based on templates for managers to announce to nurses three weeks ahead of the staffing cycle. If nurses request to swap or change shifts from the announced schedule, the software guides managers to accept or reject the changes based on staffing shortages and surpluses at each shift. The end-to-end implementation was partially used in March 2023 and has been fully deployed since April 2023.

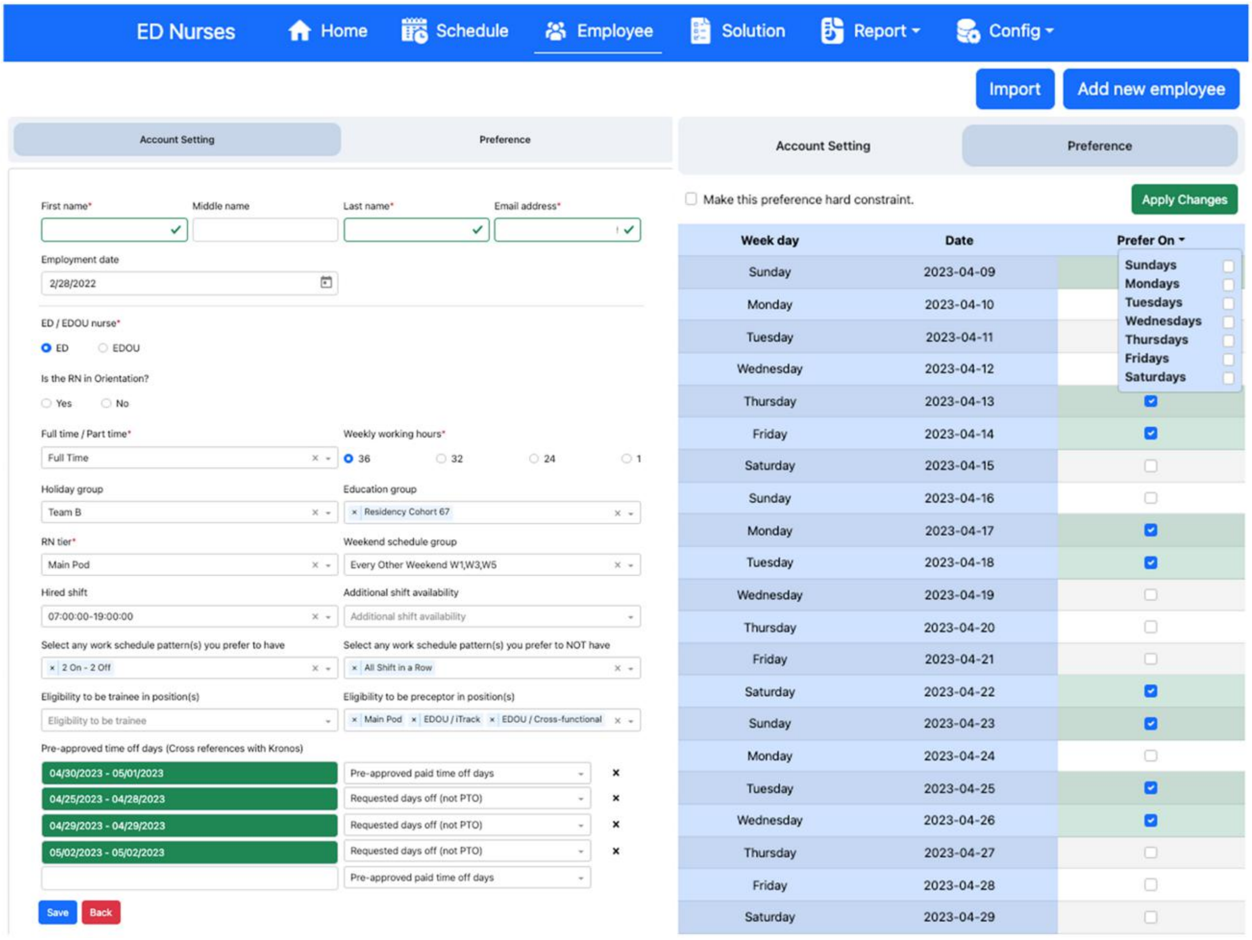
Software Illustration

In this section, we demonstrate various components of the software from input collection to output generation.

Collecting Input. The software has two sections to gather input from nurse managers and individual nurses. The Config tab enables nurse managers to set general staffing inputs: the RN Tier page specifies eligibilities for nurses of each tier to work at each position; the Groups page assigns nurses into different cohorts with corresponding unavailability dates for each cohort (e.g., holiday); the Schedule Date page defines the start and end date of the staffing cycle to be scheduled. These settings define the structure of the Employee tab shown in Figure 5. Managers (ED leadership, head nurses, and scheduling assistants) can modify the nurse roster by adding/removing employees or importing an Excel file. They can also edit each nurse's account settings and basic information. Once the forms are released, nurses can log into their accounts and enter their shift availability, shift pattern preferences, dates they prefer to work on, and any request for paid or unpaid time off. After nurses enter their preferences, managers approve or deny their request for days off, hence validating the inputs for the subsequent optimization models.

Generating Solutions. After collecting nurse input, nurse managers can use the software's output sections. In the Solutions tab, managers can set some parameters (e.g., penalties for the different objective components), solve the model, and obtain a solution (identified using a user-specified name). They can review and compare the recommended aggregate staffing levels in a table and select some modeling options such as whether to allow nurses to change shift types, to include training schedules, and to incorporate fairness considerations. The software offers additional flexibility to solve for a

Figure 5. (Color online) Software Section for Employee Information and Preference Collection

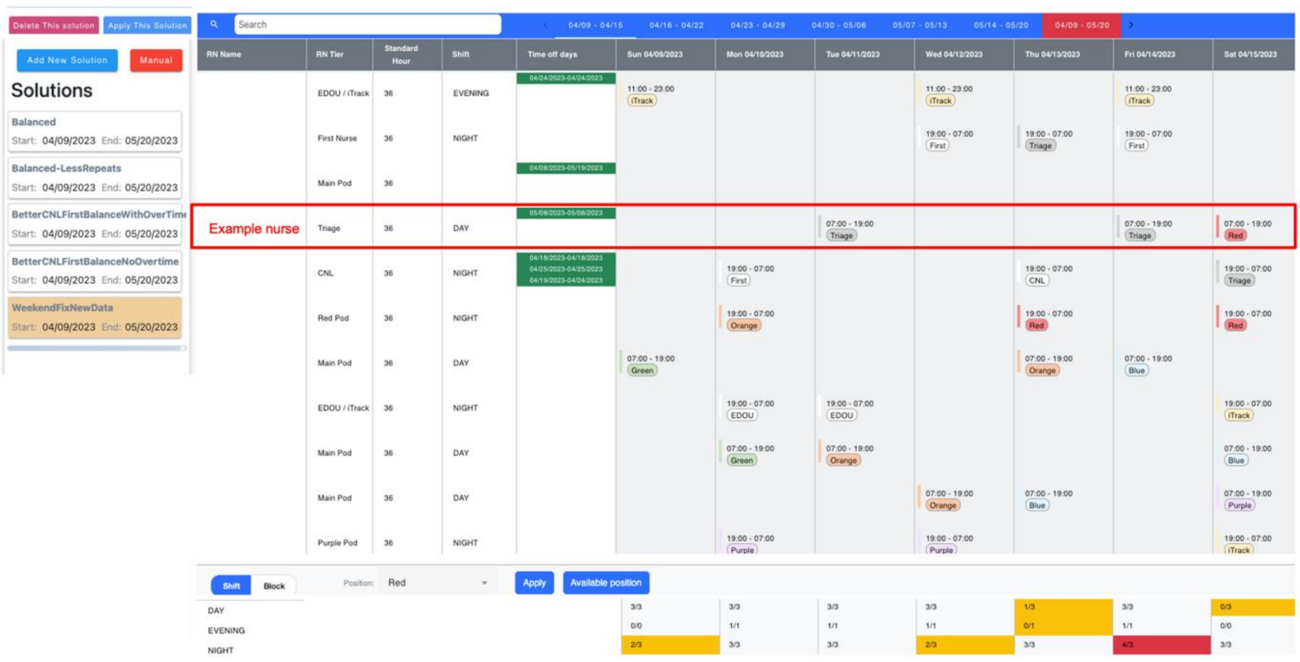


subset of nurses or tiers, such as solving for ED only or EDOU nurse only. The managers can experiment with different parameters and compare alternative aggregate or individual solutions, as shown in Figure 6. In the schedule table in Figure 6, each nurse is represented by a row containing the assigned shifts, positions (in different colors), and dates for that nurse.

Adjusting Schedules. Owing to dynamic changes in staffing needs and nurse availabilities, the support for schedule adjustments is deemed essential. As seen at the bottom of Figure 6, aggregate staffing levels from the individual schedule (first number) are presented in comparison with target levels (second number) for each shift and any subset of positions. For example, Friday, April 14, has a surplus of one nurse during the 7 p.m.–7 a.m. shift (highlighted in red), whereas Saturday, April 15, has a shortage of one nurse during the 7 a.m.–7 p.m. shift (highlighted in orange). Every week, the software also resolves the model with updated demand and alerts managers of any expected changes. This component raises managers' attention on potential shortages or surpluses across the pods, suggesting adjustments to be made. Also demonstrated in Figure 6, managers can execute any schedule changes by editing shifts, dates, positions, or nurses with simple clicks. The combination of these functionalities supports a variety of schedule adjustments: it suggests how each shortage can be filled if specific nurses can work overtime, advising managers of specific nurses that can help reduce the shortages best. Because some nurses sometimes request shift swaps owing to changes in their availability, the aggregate summary helps their managers approve these requests by visualizing whether the affected shifts are currently experiencing a surplus or shortage of nurses. Similarly, if two nurses request to swap with each other, the software automatically checks the feasibility of the switch. To cover additional shortages, managers can add shifts on per diem nurses (who do not work full-time but pick up shifts on demand anytime) to fill in the shortage shifts.

Announcing Outputs. Finally, the Schedule tab shows the determined schedule, where nurse managers have

Figure 6. (Color online) Software Section for Generating and Adjusting Alternative Individual Schedules



a holistic view, and each nurse can view their own schedule. In addition, the Report tab enables managers to download and print reports of the schedule based on the ED template, including a summary of staffing for each cycle and a daily team sheet for each date. Figure 7 includes an example team sheet on April 9, 2023, which assigns nurse names at each position during each time block for both day and night shifts of the day. Overall, the software provides decision support as well as aligns with ED operations.

Financial Benefit Estimation

The schedule optimization with its implementation translates into substantial financial benefits projected by HH. The reduction of nursing hours from Tables 3 and 4 converts into estimated cost savings in Table 5. By saving 2.6 shifts (12 hours long) daily, ED will save an average of 31.2 nurse-hours per day, 11,388 hours per year. With overtime salary being 50% above the base salary, the reduction of 1.1 overtime shifts per day cuts the cost by an additional $25 per hour. Combined with the 88 hours of manual work per six weeks that could be saved from Table 4, the hospital financial department projects that the optimization could save $728,088.10 annually in nursing costs.

Overcoming Challenges in Deployment

Deploying an optimization tool for ED nurse staffing, which impacts the lives of 200 nurses, nurse schedulers, and managers, carries numerous challenges. The deployment of the tool also involved transitioning from an offline to an online scheduling process. In this section, we discuss how we overcame these challenges, progressed with the deployment, and achieved practical impact.

Offline Strategic Adoption. The development started with offline iterations in late 2020. Researchers generated schedules offline for ED leadership to review and then incorporated feedback and constraints into the models. As described in the 2021: Accounting for Frictions to Implementation section, researchers adapted the optimization models and parameters to overcome the nurse leadership's conservatism in reducing staffing levels. With refined models, the team communicated with hospital leadership to adopt our optimized schedules. The model suggested fewer 7–7 shifts and more 11–11 shifts in response to demand patterns. However, as each nurse was preassigned to a fixed shift type upon being hired, it was difficult to swap shifts without disrupting their lives. After strategic discussions, hospital leadership decided to hire more 11 a.m.–11 p.m. shift nurses as opposed to 7–7 shift nurses to be able to realize the recommendation. The ED reports having tripled the number of 11 a.m.–11 p.m. positions (from 4.9% to 14.7% of total nurses). During this overstaffing period, the optimized schedule needed fewer nurse shifts than previously as shown in the End of 2020: Reducing Staffing Costs During Low-Demand Periods section. This presented another challenge as the

Figure 7. (Color online) Software Section for Outputting Final Schedules

	ED ASSIGNMENT RECORD: 2023-04-09 DAY SHIFT			ED ASSIGNMENT RECORD: 2023-04-09 NIGHT SHIFT	
	07:00:00-11:00:00	11:00:00-15:00:00	15:00:00-19:00:00	19:00:00-23:00:00	23:00:00-07:00:00
CNL	RN name	RN name	RN name	RN name	RN name
Triage	RN name	RN name	RN name	RN name	RN name
Triage	RN name	RN name	RN name	RN name	
Triage	RN name	RN name	RN name		
Triage		RN name	RN name		
Blue					
12, 13, 14, 15, 16 LEAD	RN name	RN name	RN name	RN name	RN name
7, 8, 9, 10, 11	RN name	RN name	RN name	RN name	RN name
17, 18, 19, 20, 21	RN name	RN name	RN name	RN name	RN name
22, 23, 24, 25, 26				RN name	RN name
Green					
38, 39, 40, 41 LEAD	RN name	RN name	RN name	RN name	RN name
27, 28, 29, 32, 32H	RN name	RN name	RN name	RN name	RN name
30, 31, 33, 34, 34H	RN name	RN name	RN name	RN name	
35, 35H, 36, 36H, 37		RN name	RN name		
Orange					
73, 74, 75, 83, 84 LEAD	RN name	RN name	RN name	RN name	RN name
62, 63, 64, 65, 66, 67	RN name	RN name	RN name	RN name	RN name
68, 69, 70, 71, 72	RN name	RN name	RN name	RN name	
76, 77, 78, 80a, 80b		RN name	RN name		
79, 81a, 81b, 82a, 82b					
Red					
Blue Functional Float (BFF)	RN name	RN name	RN name	RN name	RN name
Green Functional Float (GFF)	RN name	RN name	RN name	RN name	RN name
Orange Functional Float (OFF)	RN name	RN name	RN name	RN name	RN name
Purple					
West RN (LEAD)	RN name	RN name	RN name	RN name	RN name
West RN	RN name	RN name	RN name	RN name	
South RN		RN name	RN name		
iTrack	RN name	RN name	RN name	RN name	
iTrack	RN name	RN name	RN name		
iTrack		RN name	RN name		
EDOU	RN name	RN name	RN name	RN name	RN name
EDOU	RN name	RN name	RN name	RN name	RN name
EDOU	RN name	RN name	RN name	RN name	RN name
EDOU	RN name	RN name	RN name		
First				RN name	RN name

hospital was still required to use all nursing hours as per the contract. After evaluating several options, the ED decided to turn the saved work shifts into additional training shifts, which are counted as working hours but do not serve patient demand. In addition, scheduling fewer weekend shifts changed the ED's conventional policy to move nurses from working every other weekend to every third weekend after working for two years as a seniority incentive. Since January 2021, the ED was able to move a total of 80 nurses from every other weekend to every third weekend early (between two and seven months prior to the two-year mark), which enhances retention rewards to nurses.

Motivation from Offline to Online. Following the successful strategic adoptions at the ED, our next goal was to implement the generated individual nurse schedule for the ED to use. However, experimentation for the next staffing cycle showed that offline iterations between researchers solving models and managers providing feedback were not feasible to adopt. Managers needed to propose changes to the schedule frequently for reasons such as nurse dissatisfaction, availability changes, and manager preferences. The frequent back-and-forth communications with the research teams caused delays and were not sustainable. Conversely, manually reshuffling the machine-generated schedule was infeasible for nurse managers without optimization training. To

Table 4. Estimated Time Spent by ED Leaders on Manual Scheduling

Responsibility component	Responsible staff	Total time
Schedule build/balancing	Scheduler	12 hours
Schedule approval	Scheduler/manager	1 hour
Management of shift switches	Scheduler/manager/assistants	4 hours weekly
Orientation scheduling	Educators/managers/scheduler	3 hours
Daily team sheet building	Manager/assistants	8 hours weekly
Total hours per six-week staffing cycle		88 hours

Table 5. Projected Savings of Nursing Hours and Cost for HH ED

Component	Daily hours	Annual hours	Hourly cost ($)	Annual cost ($)
Base shifts	31.2	11,388.00	50	569,400.00
Overtime (extra)	13.2	4,818.00	25	120,450.00
Manual scheduling	2.1	764.76	50	38,238.10
Total				728,088.10

address these issues, we decided to develop our software tool, which automates the optimization part and allows nurse managers to regenerate and edit schedules in an online fashion.

Online Implementation. In 2022, we began the software development for staffing automation. The research team collaborated with the consultancy and IT teams and integrated the optimization models into the software. Given imported input data, users were able to solve the models and obtain the schedule under the Solution and Schedule tabs. However, the upcoming main challenge was to connect the software with the hospital to collect input from nurses and pass output for the ED to use. One complication arose as the entire nurse staffing at Hartford HealthCare relied on commercial software that links shift assignments to the nurse payroll system. Despite the team's attempts to integrate with the commercial software throughout 2022, the complexity of involving a third-party organization was beyond our control. After careful discussions between the team and executives at Hartford HealthCare, a decision was made to build a standalone software instead. In early 2023, we built the information collection component of the software. After frequent weekly meetings and iterations between researchers, the development team, nurse managers, and schedulers, the interface evolved to cover all ED nurse staffing functionalities. We further improved the schedule generation process, enabling dynamic schedule changes by nurse leadership to adapt to quick information changes. The ED executed more model recommendations after the long collaboration, such as swapping shift types for some nurses. The fine-tuned schedule announced on April 30, 2023, has no overtime shifts, satisfies 92% preferred date assignments, has only one undesirable shift pattern among all nurses across six weeks, and assigns nurses with desirable diversity to an average of 2.43 different positions per week. The impact of the tool's deployment can be evaluated more deeply as more data after the implementation is collected in the future. Following the pilot ED deployment at HH, the Hartford HealthCare executive team aims to extend the models and software to cover all nurse staffing in seven hospitals of the network in the future.

Conclusions

Under a collaboration between Hartford HealthCare, MIT, and Dynamic Ideas, we develop and implement optimization models to automate the nurse staffing process in the emergency department at HH. Our methodology consists of two phases to optimize each six-week staffing cycle. First, we learn an uncertainty set from patient demand data and develop a robust optimization model, solved via a cutting-plane algorithm to compute aggregate staffing levels. Next, we develop two mixed-integer optimization models to assign individual nurses to work, trainee, and preceptor shifts across the coming six-week period.

Experimental results demonstrate the versatile benefit of the first-phase aggregate model: reducing staffing costs by 5%–8% in pre–COVID-19 periods during demand was relatively low compared with nurse supply versus reducing insufficiency by 8%–25% since the COVID-19 pandemic. In addition, we analyze how the outcomes change with different model variations and parameter tradeoffs and illustrate schedule iterations and demand patterns. The second optimization phase produces individual schedules of nurses to shifts and brings significant benefits to ED nurses by reducing 17% weekend, 14% holiday, and 85% overtime shifts while increasing 5% satisfaction score, 3.6 more diverse positions, and 0.95 training shifts per day.

The models are integrated into end-to-end software that supports scheduling from staffing preparation to input collection, solution generation, and schedule output. After overcoming numerous implementation challenges, the software was deployed starting March 2023 at HH, automating and transforming the nurse scheduling at the emergency department. The implementation relieved manual scheduling burdens (78 hours per six-week period) and is projected to save $728,088.10 in annual nursing costs. The integration brought more cost-effective, sufficient, and desirable staffing into practice, benefiting various stakeholders in the hospital system.

Acknowledgments

The authors thank the Hartford HealthCare team, especially Cheryl Ficara, Barry Stein, Yolanda Johnson, and Katherine Martinez-Taveras, for support on the implementation. L. Na led the efforts of developing the model, planning and performing experiments, writing code, analyzing results, and writing the manuscript and contributed to software implementation. J. Pauphilet contributed to the design of research, modeling, and experiments, as well as manuscript writing. A. Haddad-Sisakht and L. Raison implemented the software and maintained the production.

A. Silver, P. Veronneau, and N. Vogt led the nursing team, executed deployment at the hospital, provided feedback, and contributed to the model improvement. D. Bertsimas directed the overall project, from concept and research to implementation, and edited the manuscript.

Appendix

Demand Patterns and Standard Scheduling

Demand Trend over Time. Figure A.1(a) displays the number of ED patient arrivals for each day since late 2016. Before March 2020, the ED volume was about 250–350 patient arrivals per day with small fluctuations. At the beginning of the COVID-19 pandemic, the number of ED arrivals dropped significantly and then gradually increased throughout 2021, up to a level slightly lower than before, around 200–300 arrivals per day. As the pandemic progressed through different stages, people were recovering from the COVID-19 isolation period, which brought the ED volume back to 250–350 arrivals per day while having larger fluctuations over time compared with pre–COVID-19. Given the significant disruption incurred during and since the COVID-19 outbreak, we use data starting from June 22, 2020, for our analysis, which we highlight in Figure A.1(b).

Demand Patterns and Baseline Staffing Levels. In Figure 1, we present the demand patterns from November 2 to December 13, 2020. In this period, the average number of patient stays over six weeks and the number of nurses recommended to staff each week in each pod type from Monday to Sunday (starting from 7 a.m. to 7 a.m. each day) are shown in Figure 1(a) and Figure A.2(a). For intraday patterns, we show the average number of patients over six weeks in each pod type at each hour of the day (from 0000 hours to 2300 hours) in Figure 1(b). For the main, red, and iTrack pods, we have higher demands in the afternoons and evenings, during which the variability is highest for main pods. For the purple pod, demand is more constant throughout the day with slightly higher demand at nighttime. Conventionally, every day applies the schedule shown in Figure 1(c) for the number of nurses working in each position at each hour of the day. Only the staffing level in iTrack changes within the day, with more nurses in afternoons and evenings and fewer at night. However, in the red and main pods, in which we also observe variability in demand (especially in the main pods), staffing levels remain constant throughout the day and throughout the week.

Robust Optimization Model for Aggregate Staffing

Here, we provide details on the robust optimization model for determining the aggregate staffing levels described in the ED Aggregate Staffing Optimization section.

Aggregate Model Data. Main Components. Allocating ED nurse staffing levels contains the following basic components.

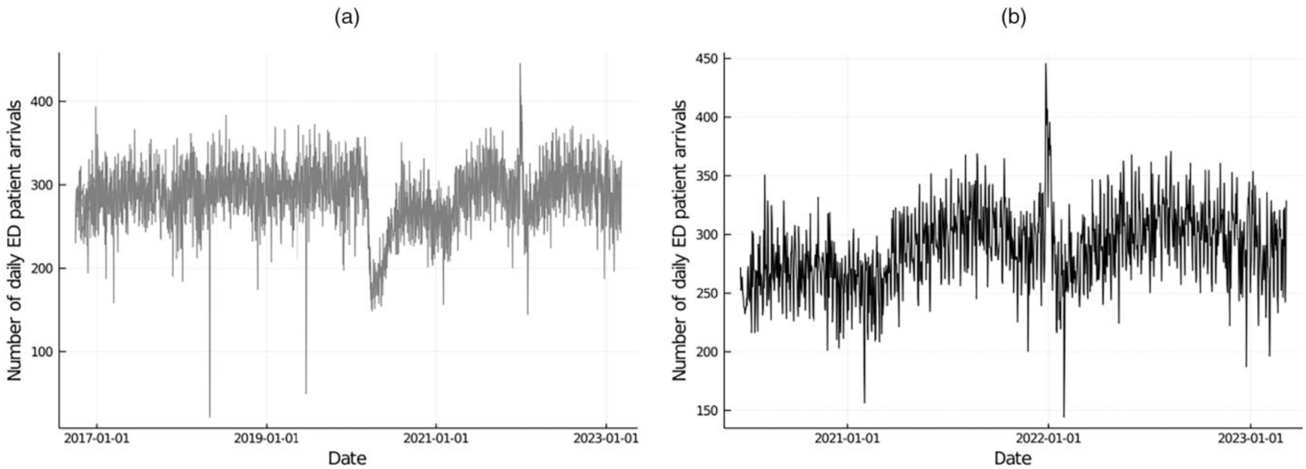
ED Positions: Nurses work among $J = 13$ positions: CNL, first nurse, resource, triage, FEP, blue, green, orange, hallway, red, purple pod, iTrack, and EDOU. As some of the pods have the same functionalities of treatment, we define $N = 5$ pod types as main pods (blue, green, orange, and hallway), red pod, purple pod, iTrack, and EDOU. As nurses can "float" between iTrack, red, and main pods within the same shift, we define $M = 3$ pod floating groups in which each group is a set of pods among which nurses can float between.

Nurses: Based on each nurse's years of experience and training qualifications, they are categorized into $Q = 9$ nurse tiers and are also classified by the hospital into $G = 2$ groups that restrict the frequency of weekend shifts to every other weekend or every third weekend.

Time: There are $I = 3$ shift types: 7 a.m.–7 p.m., 7 p.m.–7 a.m., and 11 a.m.–11 p.m. The optimization model schedules for an entire staffing cycle (six-week period) with some week-over-week regularity. We divide each cycle into $W = 6$ weeks or $D = 42$ days or $T = 1{,}008$ one-hour time periods. Because there might be some week-over-week regularity, we also introduce indices to divide a typical week of a cycle into $E = 7$ days or $S = 168$ hours.

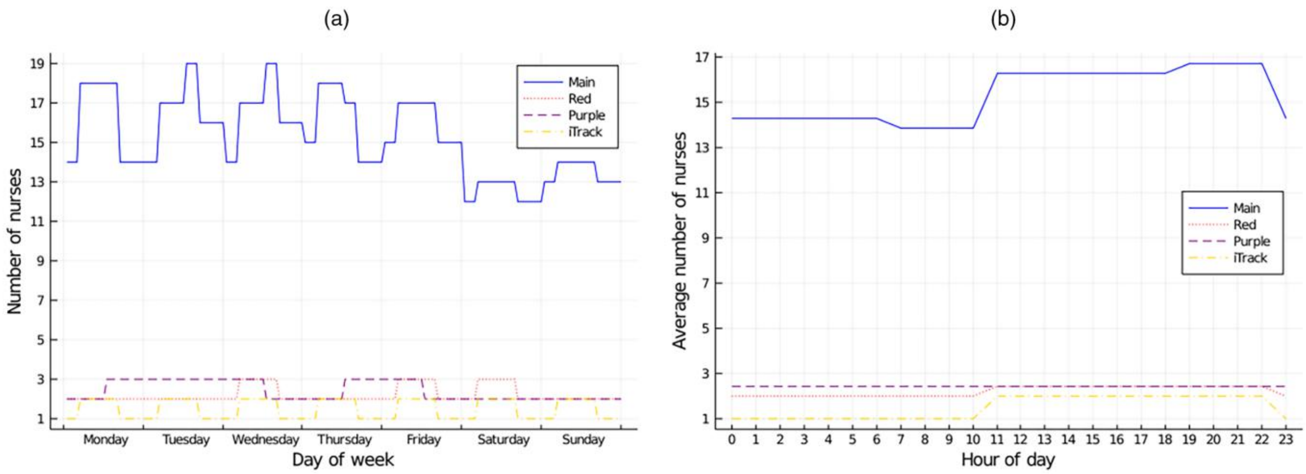
Other Sets: For each component of size J, we use the corresponding lowercase letter j to index the element, which takes values in set $[J]$ denoting $\{1, 2, \ldots, J\}$. We define other

Figure A.1. ED Historical Volume over Time



Notes. (a) From November 7, 2016, to March 6, 2023. (b) From June 22, 2020, to March 6, 2023.

Figure A.2. (Color online) Recommended Schedules for the November 2–December 13, 2020, Cycle



Notes. (a) Over a typical week. (b) Within a day (averaged).

discrete subsets needed: J_n, $J_m \subseteq [J]$: sets of pods j of type n and of floating group m, respectively; $Dn \subseteq [D]$: weekend days of the cycle; $De \subseteq [E]$: days of the first week of the cycle; D_w, $Dn_w \subseteq [D]$: days and weekend days in week w, respectively.

Demand Uncertainty Set. For a six-week cycle starting on week w^0, we construct an uncertainty set around the demand for each pod floating group m, on each week $w \in \{w^0, \ldots, w^0+5\}$, and in each hour of the week $s \in [S]$, $\tilde{h}_{mws}$ in the following way. First, we compute a nominal demand $\bar{h}_{jws}$ from the historical demand on weeks $w^0 - 10$ through $w^0 - 5$:

$$\bar{h}_{mws} := h^{\text{hist}}_{ms(w-10)}, \quad \forall w \in \{w^0, \ldots, w^0+5\}.$$

Second, we estimate the pairwise absolute differences of historical demands that are 10 weeks apart, for the weeks between $w^0 - 20$ and $w^0 - 5$:

$$\hat{\varepsilon}_{msw} = |h^{\text{hist}}_{msw} - h^{\text{hist}}_{msw-9}|, \quad \forall m \in [M], s \in [S], w \in \{w^0-11, \ldots, w^0-5\}.$$

Finally, we define the final uncertainty set as

$$\mathcal{U} := \left\{ \tilde{h}_{msw} \geq 0 \left| \begin{array}{ll} |\bar{h}_{msw} - \tilde{h}_{msw}| \leq \epsilon^1_{ms}, & \forall m \in [M], \\ & s \in [S], w \in [W] \\ \left| \displaystyle\sum_{s \in [S]} \bar{h}_{msw} - \sum_{s \in [S]} \tilde{h}_{msw} \right| \leq \epsilon^2_m, & \forall m \in [M], \\ & w \in [W] \\ \left| \displaystyle\sum_{m \in [M], s \in [S]} \bar{h}_{msw} - \sum_{m \in [M], s \in [S]} \tilde{h}_{msw} \right| \leq \epsilon^3, & \forall w \in [W]. \end{array} \right. \right\},$$

where we calibrate ϵ^1_{ms} (i.e., the bound on pairwise inter-10-week demand fluctuations) as the 80th percentile of the $\hat{\varepsilon}_{msw}$ we estimated on historical data. For each m, s, we scale ϵ_{1ms} accordingly to obtain ϵ_{2m} and ϵ_3. From large deviation bounds (Bertsimas et al. 2021, table 2), we know that ϵ^2_m and ϵ^3 should scale as $\sqrt{S}$ and $\sqrt{MS}$, respectively, so we use

$$\epsilon^2_m = \beta_\epsilon \frac{1}{\sqrt{S}} \sum_{s \in [S]} \epsilon^1_{ms}, \quad \forall m \in [M], \quad \epsilon^3 = \beta_\epsilon \frac{1}{\sqrt{MS}} \sum_{m \in [M], s \in [S]} \epsilon^1_{ms},$$

with an adjustable discount factor $\beta_\epsilon \in (0,1]$.

Other Data and Parameters. Other than the demand data, we collect the following data from ED leadership:

- Z_q, Z^g_q: Total number of nurses of tier q, and those of weekend group g
- z^{curr}_{jie}: Number of nurses in position j during shift i on day e of the week in standard schedule
- $\underline{n}_j, \overline{n}_j$: Minimum and maximum number of nurses required at position j, respectively
- $\overline{Tn_q}$: Maximum total number of shifts each nurse of tier q can work in a week
- N_U: Set of pairs (q, j) such that nurse of tier q cannot work in position j owing to qualifications
- E_{qt}: Number of nurses of tier q available to work during the one-hour period t

We also introduce σ_{idt} (or equivalently, σ_{idsw}) to indicate whether shift i on day d contains time period t.

Aggregate Optimization Formulation

Decision Variables. In the next six-week cycle, on each day of the week, we decide the following:

- $z_{qjidg} \in \mathbb{Z}^+$: Number of nurses of tier q, weekend group g working at position j on shift i, day d

To capture some aspects of our objectives to minimize insufficiency and deviation from the current schedule, we introduce two auxiliary decision variables:

- npr: The weighted sum of insufficiency in the worst case (w.r.t. the demand uncertainty set), where insufficiency is defined as the number of nurse shifts missing to satisfy the target nurse-to-patient ratios to treat the patients
- Δz_{jie}: The absolute difference in the number of scheduled nurses working at position j during shift i on day e of the week compared with the standard schedule

Objective. The objective of the model is to minimize the total number of nurse shifts scheduled while penalizing shortage for demand and changes from the standard schedule, with parameters to control tradeoffs between the three terms. Hence, our objective is to minimize a weighted combination of three terms:

$$\min \sum_{q\in[Q],j\in[J],i\in[I],d\in[D],g\in[G]} z_{qjidg} \quad \text{(Minimize number of scheduled work shifts)}$$
$$+ \mu_1 \cdot npr \quad \text{(Penalize when staffed below target ratios)}$$
$$+ \mu_2 \sum_{j\in[J],i\in[I],e\in[E]} \Delta z_{jie} \quad \text{(Penalize number of changed work shifts)},$$

where the parameters $\mu_1, \mu_2 \geq 0$ control the relative importance of each term of the objective.

Constraints. We now present the constraints. Some of these constraints correspond to the definition of the auxiliary variables and the others capture various staffing requirements. We use the notation $f(x)_+$ to denote the positive part of the function $f(x)$—that is, $f(x)_+ := \max\{f(x), 0\}$.

For defining the auxiliary variables, we impose the following constraints:

$$npr \geq \sum_{m\in[M],s\in[S],w\in[W]} \omega_w \left(\tilde{h}_{msw} - \sum_{q\in[Q],j\in J_m,i\in[I],d\in[D],g\in[G]} z_{qjidg}\sigma_{idsw}\right)_+, \quad \forall \tilde{h} \in \mathcal{U}, \tag{A.1}$$

$$\Delta z_{jie} \geq \left| z^{\text{curr}}_{jie} - \sum_{q\in[Q],d\in De,g\in[G]} z_{qjidg} \right|, \quad \forall j \in [J], i \in [I], e \in [E]. \tag{A.2}$$

Constraint (A.1) computes the difference between the nurse demand for each position/hour/week $\tilde{h}_{msw}$ and the total number of nurse shifts scheduled for this position/hour/week—the coefficients $\sigma_{idsw}\{0,1\}$ map shifts (i) and day (d) to each hour s and week w. It is a robust constraint imposed for all demand realization $\tilde{h} \in \mathcal{U}$. It also involves a weighting parameter ω_w, which allows us to put more weight on some weeks. For example, by increasing ω_w with w, we place higher penalties for the weeks estimated with the more recent data, hence emphasizing the most recent demand trend.

The schedule must adhere to a range of staffing constraints to be feasible. These include the following:

- Each position j has a minimum and maximum number of nurses to staff (staffing levels are typically fixed at logistical positions, e.g., CNL and first nurse, to be the same as before, whereas some pods can be adjusted to demand): $\underline{n}_j \leq \sum_{q\in[Q],i\in[I],d\in[D],g\in[G]} z_{qjidg} \cdot \sigma_{idt} \leq \overline{n}_j, \quad \forall j \in [J], t \in [T]$.
- Certain units are not eligible based on nurse tiers: $z_{qjidg} = 0, \quad \forall (q,j) \in N_U, i \in [I], d \in [D], g \in [G]$.
- Staffing levels are kept consistent from week to week for a stable schedule:

$$\sum_{q\in[Q],g\in[G]} z_{q,j,i,k+E(w-1),g} = \sum_{q\in[Q],g\in[G]} z_{q,j,i,k+E(W-1),g}, \quad \forall k \in [E], w \in [W-1], i \in [I], j \in [J]. \tag{A.3}$$

ED leadership indicates a preference to consider only the constraints described here when deciding on staffing levels, based on demand only. However, they have the option to also consider nurse availability and supply information in this stage of scheduling by including additional constraints in the model:

- Assignments are capped by the number of available nurses of each tier during each period: $\sum_{j\in[J],i\in[I],d\in[D],g\in[G]} z_{qjidg} \cdot \sigma_{idt} \leq E_{qt}, \quad \forall q \in [Q], t \in [T]$.
- Total number of weekly working hours: $\sum_{d\in D_w,j\in[J],i\in[I],g\in[G]} z_{qjidg} \leq \overline{Tn_q} Z_q, \quad \forall q \in [Q], w \in [W]$
- Weekend shifts on consecutive weeks for nurses working every other weekend (and analogous constraints for every third weekend): $\sum_{d\in Dn_w \cup Dn_{w+1},j\in[J],i\in[I],g\in[G]} z_{qjidg} \leq 2Z^1_q$, $\forall q \in [Q], w \in [W-1]$

Solving the Robust Optimization Model

We now describe our method and implementation of solving the robust optimization problem. We use a cutting-plane approach shown in Algorithm A.1. At each iteration κ, we solve a mixed-integer linear optimization problem (a master problem) similar to the one described earlier, except that the uncertainty set $\mathcal{U}$ in Constraint (A.1) is replaced by a finite subset $\mathcal{U}^\kappa$. To solve the master problem, we linearize the positive part in Constraint (A.1) and the absolute values in Constraint (A.2) via additional auxiliary variables and linear constraints. At each iteration, we augment $\mathcal{U}^\kappa$ with the worst-case demand scenario h^κ and solve the master problem up to 0.01 optimality within a 20-minute time limit. We implement the cutting-plane algorithm using lazy callbacks to accelerate the model compilation and solving process. As termination criteria, we use a suboptimality gap target $\eta = 0.1$ and a maximum number of iterations $\kappa^{\max} = 20$.

Algorithm A.1 (Cutting-Plane Algorithm)

1: Given the nominal demand $h^{\text{hist}}: \mathcal{U}^0 \leftarrow \{h^{\text{hist}}\}$, $\kappa \leftarrow 1$, and tolerance η, initiate the master problem
2: **repeat**
3: Solve the master problem and obtain a solution z^κ
4: $npr_\kappa(\tilde{h}, z^\kappa) \leftarrow \sum_{m\in[M],s\in[S],w\in[W]} \omega_w (\tilde{h}_{msw} - \sum_{q\in[Q],j\in J_m,i\in[I],d\in[D],g\in[G]} z^\kappa_{qjidg} \sigma_{idsw})_+$
5: $h^\kappa \leftarrow \arg\max_{\tilde{h}\in\mathcal{U}} npr_\kappa(\tilde{h}, z^\kappa)$
▷ Maximize insufficiency w.r.t. demand uncertainty set
6: $\mathcal{U}^\kappa \leftarrow \{h^{\text{hist}}, h^1, \ldots, h^\kappa\}$
7: $\kappa \leftarrow \kappa + 1$
8: **until** $(npr^\kappa - npr^{\kappa-1})/npr^{\kappa-1} < \eta$ or $\kappa = \kappa^{\max}$
▷ Reach violation gap or maximum iterations

We fine-tune a combination of parameters $(\beta_\epsilon, \mu_1, \mu_2, \omega_w)$ to select the best-performing values for the model. To solve the deterministic model version, we solve a single master problem using the optimization model defined in the Optimization Model Overview section, but with $\mathcal{U}$ replaced by the set of historical demand only $\{h^{\text{hist}}\}$ in Constraint (A.1), which is equivalent to setting $\epsilon^1 = \epsilon^2 = \epsilon^3 = 0$ in $\mathcal{U}$.

Model Variants for Informing Strategic Decisions

Stable Staffing Levels. Motivated by the within-week variability of demand, we allow the staffing levels to vary by day of the week—but remaining the same week over week as per Constraint (A.3). To evaluate the benefits of this additional flexibility, we compare it with two alternatives. *Daily stability*: Staffing levels can be forced to be the same every day by making weekly Constraints (A.3) daily. This option is easier for the leadership to manage but might result in unnecessary overstaffing and understaffing on some days. It corresponds to the historical practice at HH. *No stability*: Staffing levels can be allowed to vary every day by excluding Constraint (A.3). This option provides even more flexibility to fit demand patterns that may differ from one week to another but is more at risk for overfitting the data. We compare the operational costs and benefits of these options.

Change Shift Designs. ED leadership also considers changing or adding shift types. Shift changes would be disruptive and require nurses to change their lifestyles to accommodate for the new working hours. However, new shifts could lead to more cost-effective staffing and a better matching of the demand patterns. With many potential shift type options available, it is essential to determine which ones could yield the greatest improvement. We support this investigation by introducing binary variables $y_i \in \{0,1\}$ that indicate whether a shift type i is introduced. We have the set of existing shift types I_{exist} and consider a feasible set of potential shift types $I_{\text{feasible}} \supseteq I_{\text{exist}}$ with σ_{idt} for each $i \in I_{\text{feasible}}$. We add the following constraints:

- A shift type is generated if and only if used in any shifts: $\sum_{q\in[Q],j\in[J],d\in[D],g\in[G]} z_{qjidg} \le M_y \cdot y_i, \quad \forall i \in [I]$ with a constant M_y.
- The number of shift types is bounded with a parameter $Y_{\max}$: $\sum_{i\in I_{\text{feasible}}} y_i \le Y_{\max}$.
- If considering only new shift types without changing the existing shift types, then all current shift types are kept: $y_i = 1, \quad \forall i \in I_{exist}$.

By solving different variations and comparing their objective values, we can identify the most valuable shift type candidates.

Mixed-Integer Optimization Models for Individual Scheduling

We provide details on the individual shift and preceptor scheduling optimization models from the ED Individual Scheduling Optimization section.

Individual-Level Data

We define additional indices, sets, and data inputs to the Individual Model Data section.

Additional Components. We consider each individual nurse ℓ (with $L \approx 200$ total nurses), with L_q denoting the set of nurses of tier q. There are multiple ways to measure a nurse's dissatisfaction. We identify six possible criteria: the number of shifts on dates supposed to be off, unassigned weekends turned on, overtime shifts, and unpreferred shift types, dates, or patterns. We also have three criteria for satisfaction: the number of preferred shift dates, patterns ($U = 3$ on/off patterns for assigned shifts), and unique positions a nurse is assigned to (as a measure of job diversity). We index these criteria via a subscript $o \in [O]$ ($O = 9$). For the ED, J^{slack} represents set of positions j where a shortage or surplus of at most one is allowed.

Data Input. One of the inputs is the output aggregate levels z^{agg}_{jid}, representing the number of nurses to schedule for position j during shift i on day d from the first phase solution $z^\star_{qjidg}$:

$$z^{\text{agg}}_{jid} = \sum_{q\in[Q],g\in[G]} z^\star_{qjidg}, \quad \forall j \in [J], i \in [I], d \in [D].$$

The assignment per nurse tier q and weekend group g in the first phase is included to ensure a feasible assignment, but it is subject to change during the second phase. The other type of input is a range of individual preferences and availabilities for each nurse: $I^{\text{pref}}_{\ell i}, I^{\text{unpref}}_{\ell i}, I^{\text{feas}}_{\ell i}$, Whether nurse ℓ prefers (typically their current shift type), does not prefer, and has the feasibility to work at shift type i, respectively. $F_{\ell w}$ is the maximum number of shifts for nurse ℓ on week w, obtained by subtracting PTO and education time of each week from the total number of weekly shifts; $k_{\ell w}$ is whether nurse ℓ is assigned to work on the weekend of week w; $P^{\text{pref}}_{ld}, P^{\text{unpref}}_{ld}, P^{\text{off}}_{ld}, P^{\text{avail}}_{ld}$ are whether nurse ℓ prefers, does not prefer, is off on, and is available to work on day d, respectively; $A^{\text{pref}}_{\ell u}, A^{\text{unpref}}_{\ell u}$ are whether nurse ℓ prefers to have and not to have the work pattern u, respectively; $r^{\text{feas}}_{\ell j}, p^{\text{feas}}_{\ell j}$ are whether nurse ℓ is eligible to be a trainee and a preceptor (training the trainees), respectively, at position j.

Individual Shift Scheduling Model

To generate an individual schedule that prioritizes various staff preferences and training opportunities given these inputs, we develop a mixed-integer linear problem that comprises the following components.

Decision Variables. The core decision variables track nurse assignments.

- $b_{\ell jid}, r_{\ell jid} \in \{0,1\}$: Whether nurse ℓ works and, respectively, trains at position j during shift i on day d
- $s_{\ell i} \in \{0,1\}$: Whether nurse ℓ is assigned to shift type i

Objective. We introduce main auxiliary variables to track different terms of the objective.

- $c_{\ell o} \in \mathbb{R}$: Computing nurse ℓ's dissatisfaction penalty score according to the criterion o
- $z^-_{jid}, z^+_{jid} \in [0,1]$: Whether there is one shortage or surplus at position j during shift i on day d, respectively
- $f^-_\ell \in \mathbb{R}^+$: Number of unassigned shifts for nurse ℓ

Our objective function is to minimize a weighted combination of four terms:

$$\begin{aligned}
\min \quad & \mu_3 \sum_{\ell\in[L],\, o\in[O]} \lambda_o c_{\ell o} && \text{(Minimize total weighted dissatisfaction score)}\\
& + \mu_4 \sum_{j\in[J],\, i\in[I],\, d\in[D]} w^{\text{shortage}}_j z^-_{jid} && \text{(Penalize shortage to aggregate staffing levels)}\\
& + \mu_5 \left(\sum_{j\in[J],\, i\in[I],\, d\in[D]} z^+_{jid} + \sum_{\ell\in[L]} f^-_\ell \right) && \text{(Penalize unassigned nurse shifts and surpluses)}\\
& - \mu_6 \sum_{\ell\in[L],\, j\in[J],\, i\in[I],\, d\in[D]} r_{\ell jid} && \text{(Reward total training shifts assigned),}
\end{aligned}$$

with parameters

- $\lambda_1, \ldots \lambda_5, \lambda_6 > 0, \lambda_7, \lambda_8, \lambda_9 < 0$ as weights for each penalty score metric,
- $\mu_3, \mu_4, \mu_5, \mu_6 > 0$ to control the trade-off between the objective terms, and

- $w_j^{\text{shortage}} \geq 0$ as penalization weights for shortage at position j.

To optimize computational efficiency, we implement all variables as sparse matrices and tensors under conditions as follows. For instance, binary variables $b_{\ell jid}$ are only needed if the values are allowed to be one and thus are only defined for indices ℓ, j, i, d such that nurse ℓ is eligible to work at position j and is available during shift i on day d. Such sparse indexing naturally incorporates some eligibility, availability, and feasibility constraints.

Constraints. To facilitate the model constraints, we define additional auxiliary binary variables on whether nurse ℓ works an overtime shift on week w ($f^+_{\ell w}$); works on a weekend of week w supposed to be off ($k^+_{\ell w}$); has at least one shift at position j during the staffing cycle ($v_{\ell j}$); and starts work pattern u on day d ($a_{\ell ud}$). We then impose associated constraints to bound the objective function:

- Computes individual nurse penalty score for each $l \in [L]$ counting the number of
  - Shifts on dates supposed to be off: $c_{\ell 1} = \sum_{d\in[D], j\in[J], i\in[I]} P^{\text{off}}_{\ell d}(b_{\ell jid} + r_{\ell jid})$
  - Unassigned weekends turned on: $c_{\ell 2} = \sum_{w\in[W]} k^+_{\ell w}$, where $\sum_{d\in Dn_w, j\in[J], i\in[I]}(b_{\ell jid} + r_{\ell jid}) \leq 2(k_{\ell w} + k^+_{\ell w}), \forall \ell \in [L], w \in [W]$
  - Overtime shifts: $c_{\ell 3} = \sum_{w\in[W]} f^+_{\ell w}$, where $\sum_{d\in D_w, j\in[J], i\in[I]}(b_{\ell jid} + r_{\ell jid}) \leq Fw_{\ell w} + f^+_{\ell w}, \quad \forall \ell \in [L], w \in [W]$
  - Unpreferred shift types: $c_{\ell 4} = \sum_{d\in[D], j\in[J], i\in[I]} I^{\text{unpref}}_{\ell i} b_{\ell jid}$
  - Unpreferred dates: $c_{\ell 5} = \sum_{d\in[D], j\in[J], i\in[I]} P^{\text{unpref}}_{\ell d}(b_{\ell jid} + r_{\ell jid})$
  - Preferred dates: $c_{\ell 6} = \sum_{d\in[D], j\in[J], i\in[I]} P^{\text{pref}}_{\ell d}(b_{\ell jid} + r_{\ell jid})$
  - Different positions assigned: $c_{\ell 7} = \sum_{j\in[J]} v_{\ell j}$, where $v_{\ell j} \leq \sum_{d\in[D], i\in[I]} b_{\ell jid}, \quad \forall \ell \in [L], j \in [J]$
  - Unpreferred patterns: $c_{\ell 8} = \sum_{d\in[D-2], u\in[2]} A^{\text{unpref}}_{\ell u} a_{\ell ud} + \sum_{d\in[D-3]} A^{\text{unpref}}_{\ell 3} a_{\ell 3d}$
  - Preferred patterns: $c_{\ell 9} = \sum_{d\in[D-2], u\in[2]} A^{\text{pref}}_{\ell u} a_{\ell ud} + \sum_{d\in[D-3]} A^{\text{pref}}_{\ell 3} a_{\ell 3d}$, where all shifts in a row are tracked by $\sum_{j\in[J], i\in[I], d'\in\{d,d+1,d+2\}}(b_{\ell,j,i,d'} + r_{\ell,j,i,d'}) - 2 \leq 3a_{\ell 1d} \leq \sum_{j\in[J], i\in[I], d'\in\{d,d+1,d+2\}}(b_{\ell,j,i,d'} + r_{\ell,j,i,d'}), \forall \ell \in [L], d \in [D-2]$, and other patterns by analogous constraints
- The individual nurse schedule matches aggregate staffing levels with shortages and surpluses: $z^{\text{agg}}_{jid} = z^-_{jid} - z^+_{jid} + \sum_{\ell\in[L]} b_{\ell jid}, \quad \forall j \in J^{\text{slack}}, i \in [I], d \in [D]$, where some positions do not allow shortages or surpluses: $z^-_{jid} = z^+_{jid} = 0, \quad j \in J \setminus J^{\text{slack}}, i \in [I], d \in [D]$.
- The number of unassigned shifts for each nurse is the number of maximum shifts minus assigned shifts: $f^-_\ell \geq \sum_{w\in[W]} F_{\ell w} - \sum_{j\in[J], i\in[I], d\in[D]}(b_{\ell jid} + r_{\ell jid})$.

Besides these constraints bounding the objective, the schedule is enforced to satisfy a variety of feasibility constraints, such as

- Eligibility and availability:
  - Nurse tier eligibility in positions: $b_{\ell jid} = 0, \quad \forall q \in [Q], (q,j) \in N_U, \ell \in L_q, i \in [I], d \in [D]$
  - Shift only at one location on available dates: $\sum_{j\in[J]}(b_{\ell jid} + r_{\ell jid}) \leq P^{\text{avail}}_{\ell d}, \quad \forall \ell \in [L], i \in [I], d \in [D]$
- Shift type:
  - Assigned at most one shift type over the six-week period: $\sum_{i\in[I]} s_{\ell i} \leq 1, \quad \forall \ell \in [L]$
  - Each nurse can be only assigned a feasible shift type: $s_{\ell i} \leq I^{\text{feas}}_{\ell i}, \quad \forall \ell \in [L], i \in [I]$
  - Work on assigned shift type: $\sum_{d\in[D], j\in[J]}(b_{\ell jid} + s_{\ell jid}) \leq M_s \cdot s_{\ell i}, \quad \forall \ell \in [L], i \in [I]$ with $M_s = 2D$
- Training:
  - Trainee eligibility: $r_{\ell jid} \leq r^{\text{feas}}_{\ell j}, \quad \forall \ell \in [L], j \in [J], i \in [I], d \in [D]$
  - At most three shifts per training: $\sum_{i\in[I], d\in[D]} r_{\ell jid} \leq 3, \forall \ell \in [L], j \in [J]$
  - Capped by number of eligible preceptors: $\sum_{\ell\in[L]} r_{\ell jid} \leq \sum_{\ell\in[L]} b_{\ell jid}\, p^{\text{feas}}_{\ell j}, \quad j \in [J], i \in [I], d \in [D]$
- To ensure fairness among nurses, each nurse's penalty score cannot exceed a bound $c^{\max}$:

$$\sum_{o\in[O]} c_{\ell o} \leq c^{\max}, \quad \ell \in [L]. \tag{A.4}$$

Individual Preceptor Scheduling Model

The preceptor scheduling model from the Integer Optimization Model on Preceptor Scheduling section is as follows.

Decision Variables. We introduce decision variables:

- $p_{\ell jid} \in \{0,1\}$: Whether nurse ℓ serves as a preceptor at position j during shift i on day d
- $p^{\max} \in \mathbb{R}$: Tracking the maximum number of preceptor shifts assigned among all nurses

Objective. We set the objective function to

$$\min\ p^{\max}\ \text{(maximum number of preceptor shifts assigned to each nurse).}$$

Constraints. The model is subject to several constraints, where $b^\star_{\ell jid}$ and $r^\star_{\ell jid}$ denote the optimal solution from solving the model defined in the Individual Shift Scheduling Model section.

- $p^{\max}$ is at least the number of preceptor shifts for each nurse: $p^{\max} \geq \sum_{j\in[J], i\in[I], d\in[D]} p_{\ell jid}, \quad \forall \ell \in [L]$.
- Eligibility to train the trainees at each position: $p_{\ell jid} \leq p^{\text{feas}}_{\ell j}, \quad \forall \ell \in [L], j \in [J], i \in [I], d \in [D]$
- Preceptor and work shifts in tandem at the same position: $p_{\ell jid} \leq b^\star_{\ell jid}, \ \forall \ell \in [L], j \in [J], i \in [I], d \in [D]$.
- Each trainee shift is covered by a preceptor: $\sum_{\ell\in[L]} r^\star_{\ell jid} \leq \sum_{\ell\in[L]} p_{\ell jid}, \quad \forall j \in [J], i \in [I], d \in [D]$.

Supplementary Experimental Results

In this section, we provide additional experimental results to the Results section on each of the three periods.

Supplementary Results During Low-Demand Periods

We augment the End of 2020: Reducing Staffing Costs During Low-Demand Periods section on data from October 26 to December 20, 2020, including the illustration of schedule patterns, additional information on different approaches, and elaboration on metric tradeoffs.

Patterns of Demand vs. Staffing. Figure A.2 provides a more detailed illustration of the recommended schedule for November 2 through December 13, 2020, obtained from the Robust (a) approach. A comparison with demand patterns from Figure 1 shows that our recommendation matches staffing with demand patterns and results in a more cost-effective schedule. For the main pods and the purple pod, as patient demands tend to be higher on weekdays than weekends, we staff more nurses on weekdays than on weekends. For the red pod, we have a particularly higher demand on Friday, when we also staff more nurses. Both demand and staffing for iTrack have less variability throughout days of the week. The recommended schedule shown in Figure A.2(b) is consistent with the demand patterns by having smoother staffing in the purple pod and more staffing during 11 a.m.–11 p.m. for other pods. Although the schedule is generated with information only prior to the period, it can capture most of the day-of-week and hour-of-day demand patterns in the period prospectively, which justifies the robustness of our approach. By matching staffing with demand patterns, we reduce the staffing cost by 7.40%.

Tradeoff Between Cost and Insufficiency. There exists a tradeoff between cost and insufficiency, as staffing more nurses facilitates more sufficiency with increased cost and vice versa. We illustrate the flexibility of the optimization model to control such tradeoffs by varying parameters μ_1 and β_ϵ in Figure A.3, where scatter points from each approach represent the daily average insufficiency and cost of schedules generated with different parameters. This earlier version of the model did not include the objective term with μ_2 or the parameter ω_w. The current schedule has a daily average cost of 56 and an insufficiency of 0.17. Retrospectively, given perfect information, schedules optimized with different parameters range from having 50.43 cost and 0.01 insufficiency to having 48.57 cost and 0.06 insufficiency. Prospectively with the data-driven approach, the nonrobust schedules lead to a reduction of cost to 49.43–50.90 with the tradeoff of higher insufficiency between 0.26 and 0.20. The robust schedules (a) incorporate uncertain demand deviations to trade some cost reduction for more sufficiency, ranging from giving 50.38 cost and 0.17 insufficiency to 53.24 cost and 0.08 insufficiency.

Supplementary Results During Iteration Periods

Alternative Individual Scheduling Results. In addition to the results shown in the 2021: Accounting for Frictions to Implementation section, we solve and compare four schedules with alternative settings and parameters in Table A.1. Compared with the solution with default parameters, a variation with increased weights on overtime shifts reduces overtime with a trade-off of less training. On top of the variation, allowing changes in nurses' shift types significantly reduces overtime shifts by changing four nurses' shift types. Alternatively, assigning two EDOU nurses to float to work at ED can bring several more overtime reductions. After reviewing with ED leadership, we decided to use the schedule with shift changes because of its best overall metrics.

Figure A.3. (Color online) Cost-Insufficiency Tradeoffs

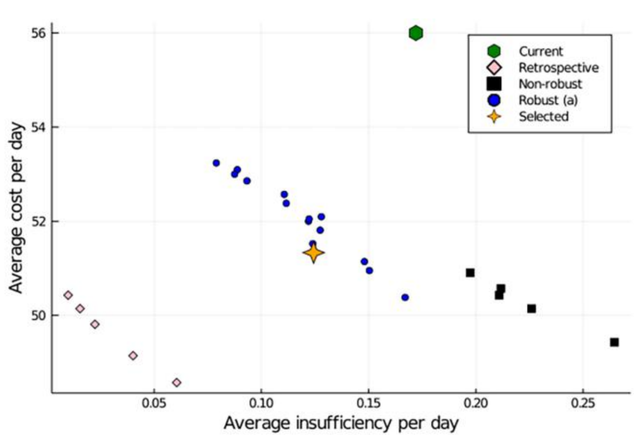


Table A.1. Comparison of Alternative RN Schedules

Variation	Default	Less overtime	Shift change	EDOU float
Overtime shifts	72	52	8	48
Training shifts	72	60	40	60
Weekends turned on	11	11	6	11
Holidays turned on	2	2	2	2
RN with shift type changed	0	0	4	0
EDOU RN relocated to ED	0	0	0	2

Supplementary Results During High-Demand Periods

We provide additional results to the 2022: Reducing Insufficiency During High-Demand Periods section.

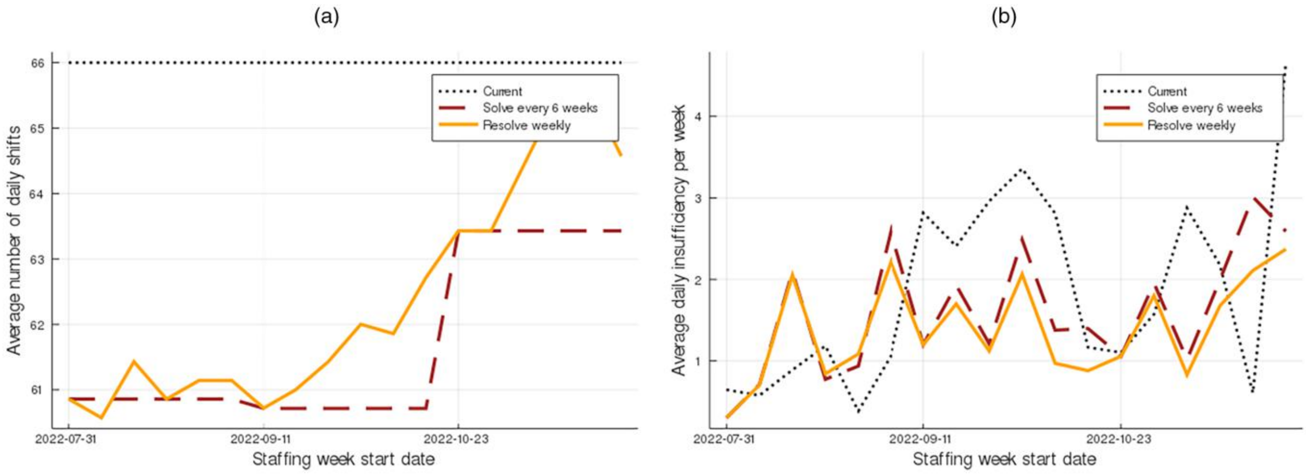
Parameter Tuning Details. We first tune the two main parameters, with μ_1 drawn from the list {0, 0.2, 0.3, 0.4, 0.6, 0.8, 1} and β_ϵ from {0.1, 0.15, 0.2, 0.3, 0.4, 0.5, 0.6} with fixed μ_2 and ω_w. With $\mu_1 = 0.2$ and $\beta_\epsilon = 0.3$ selected and fixed, we then vary μ_2 and ω_w. We use an auxiliary parameter $\omega \in [0, 1]$ and let $\omega_1, \ldots, \omega_6$ to be uniformly distributed between $1 - \omega$ and $1 + \omega$. We vary μ_2 from {0, 0,1, 0.2, 0.3, 0.4, 0.5} and vary ω from {0, 0.1, 0.2, 0.25, 0.3, 0.4, 0.5}. We finally select $\mu_2 = 0.2$ and $\omega = 0.3$ together with $\mu_1 = 0.2$ and $\beta_\epsilon = 0.3$.

Adjustment from Resolving Aggregate Model Weekly. We consider another variation to solve the aggregate model every week. We conduct experiments on three adjacent six-week staffing cycles between July 31 and December 3, 2022. We compare three alternatives:

1. Current schedule
2. The model is solved every six weeks and generates a schedule for each six-week staffing cycle. This requires solving the model three times for the three staffing cycles.
3. For each staffing cycle, the model is resolved every week using the training data one week later and update the schedule for the remaining weeks of the cycle. This involves solving the model 18 times for the 18-week period.

For each week, we compute the average number of daily shifts and average daily insufficiency during that week's schedule and plot the three variations during the time period in Figure A.4. Resolving the model weekly changes in staffing levels from week to week and leads to improved insufficiency in most weeks. We summarize the average daily metrics among the 18 weeks for the three alternatives in Table A.2. On average, resolving weekly changes the staffing levels by 2.08 shifts from week to week. This would require the nurse leadership to ask several nurses to request shift changes from their prearranged schedule. Solving the model every six weeks (respectively, every week) on average changes the staffing levels from the current schedule by 4.71 (respectively, 4.26) shifts. In return, resolving weekly can reduce the shortage of nurse shifts to meet target

Figure A.4. (Color online) Change from Resolving Weekly During Three Staffing Cycles



Notes. (a) Daily number of shifts. (b) Daily insufficiency.

Table A.2. Average Daily Results and Change (δ) from Current Schedules on Three Staffing Cycles During July 31 Through December 3, 2022

Schedule approach	Cost [δ]	Insufficiency [δ]	δz from current	δz by week
Current	66	1.85	0	0
Solve every six weeks	61.67 [−6.57%]	1.59 [−13.86%]	4.71 [7.14%]	0
Resolve weekly	62.37 [−5.51%]	1.39 [−24.83%]	4.26 [6.46%]	2.08 [3.33%]

nurse-patient ratios by 24.83% from the current schedule, which is larger than the 13.86% reduction created by solving the model every six weeks. The benefit of reducing insufficiency is particularly useful for this period when the ED is extremely understaffed and is achieved with 5.51% (respectively, 6.57%) less staffing cost from solving the model every six weeks (respectively, every week).

References

Ahmed MA, Alkhamis TM (2009) Simulation optimization for an emergency department healthcare unit in Kuwait. *Eur. J. Oper. Res.* 198(3):936–942.

Aiken LH, Clarke SP, Sloane DM, Sochalski J, Silber JH (2002) Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. *JAMA* 288(16):1987–1993.

Aiken LH, Sloane DM, Bruyneel L, Van den Heede K, Griffiths P, Busse R, Diomidous M, et al. (2014) Nurse staffing and education and hospital mortality in nine European countries: A retrospective observational study. *Lancet* 383(9931):1824–1830.

Anderson M, Bodur M, Rathwell S, Sarhangian V (2022) Optimization helps scheduling nursing staff at the long-term care homes of the city of Toronto. *INFORMS J. Appl. Analytics* 53(2):133–154.

Ang BY, Lam SWS, Pasupathy Y, Ong MEH (2018) Nurse workforce scheduling in the emergency department: A sequential decision support system considering multiple objectives. *J. Nursing Management* 26(4):432–441.

Bernstein SL, Aronsky D, Duseja R, Epstein S, Handel D, Hwang U, McCarthy M, et al. (2009) The effect of emergency department crowding on clinically oriented outcomes. *Acad. Emergency Medicine* 16(1):1–10.

Bertsimas D, den Hertog D, Pauphilet J (2021) Probabilistic guarantees in robust optimization. *SIAM J. Optim.* 31(4):2893–2920.

Bezanson J, Edelman A, Karpinski S, Shah VB (2017) Julia: A fresh approach to numerical computing. *SIAM Rev.* 59(1):65–98.

Brucker P, Qu R, Burke E (2011) Personnel scheduling: Models and complexity. *Eur. J. Oper. Res.* 210(3):467–473.

Chan CW, Huang M, Sarhangian V (2021) Dynamic server assignment in multiclass queues with shifts, with applications to nurse staffing in emergency departments. *Oper. Res.* 69(6):1936–1959.

Chen TL, Wang CC (2016) Multi-objective simulation optimization for medical capacity allocation in emergency department. *J. Simulations* 10(1):50–68.

Clark A, Moule P, Topping A, Serpell M (2015) Rescheduling nursing shifts: Scoping the challenge and examining the potential of mathematical model based tools. *J. Nursing Management* 23(4): 411–420.

Draeger MA (1992) An emergency department simulation model used to evaluate alternative nurse staffing and patient population scenarios. *Proc. 24th Conf. Winter Simulation* (ACM, New York), 1057–1064.

Dunning I, Huchette J, Lubin M (2017) Jump: A modeling language for mathematical optimization. *SIAM Rev.* 59(2):295–320.

Elder E, Johnston AN, Crilly J (2015) Systematic review of three key strategies designed to improve patient flow through the emergency department. *Emergency Medicine Australasia* 27(5): 394–404.

Gurobi Optimization LLC (2023) Gurobi optimizer reference manual. Accessed April 15, 2023, https://www.gurobi.com.

Hamid M, Tavakkoli-Moghaddam R, Golpaygani F, Vahedi-Nouri B (2020) A multi-objective model for a nurse scheduling problem by emphasizing human factors. *Proc. Institute Mechanical Engrg. H* 234(2):179–199.

Hu Y, Chan CW, Dong J (2024) Prediction-driven surge planning with application in the emergency department. *Management Sci.*, ePub ahead of print May 24, https://doi.org/10.1287/mnsc.2021.02781.

Kellogg DL, Walczak S (2007) Nurse scheduling: From academia to implementation or not? *INFORMS J. Appl. Analytics* 37(4):309–399.

Kim SH, Whitt W (2014) Are call center and hospital arrivals well modeled by nonhomogeneous Poisson processes? *Manufacturing Service Oper. Management* 16(3):464–480.

Leineweber C, Chungkham HS, Lindqvist R, Westerlund H, Runesdotter S, Alenius LS, Tishelman C, et al. (2016) Nurses' practice environment and satisfaction with schedule flexibility is related to intention to leave due to dissatisfaction: A multi-country, multilevel study. *Internat. J. Nursing Stud.* 58:47–58.

Lim G, Mobasher A (2011) Robust nurse scheduling problem. *Proc. Industrial Engrg. Res. Conf.* (IISE, Peachtree Corners, GA).

Maman S (2009) Uncertainty in the demand for service: The case of call centers and emergency departments. PhD thesis, Technion–Israel Institute of Technology, Haifa.

Mohammadian M, Babaei M, Amin Jarrahi M, Anjomrouz E (2019) Scheduling nurse shifts using goal programming based on nurse preferences: A case study in an emergency department. *Internat. J. Engrg.* 32(7):954–963.

Needleman J, Buerhaus P, Mattke S, Stewart M, Zelevinsky K (2002) Nurse-staffing levels and the quality of care in hospitals. *New England J. Medicine* 346:1715–1722.

Park CSY, Kabak M, Kim H, Lee S, Cummings GG (2022) No more unimplementable nurse workforce planning. *Contemporary Nurse* 58(2–3):237–247.

Peters M (2023) Time to solve persistent, pernicious and widespread nursing workforce shortages. *Internat. Nursing Rev.* 70(2):247–253.

Ramsey Z, Palter JS, Hardwick J, Moskoff J, Christian EL, Bailitz J (2018) Decreased nursing staffing adversely affects emergency department throughput metrics. *West. J. Emergency Medicine* 19(3):496–500.

Rerkjirattikal P, Singhaphandu R, Huynh VN, Olapiriyakul S (2022) Job-satisfaction enhancement in nurse scheduling: A case of hospital emergency department in Thailand. *Proc. 9th Internat. Sympos. Integrated Uncertainty Knowledge Model. Decision Making* (Springer, Berlin), 143–154.

Saghafian S, Austin G, Traub SJ (2015) Operations research/management contributions to emergency department patient flow optimization: Review and research prospects. *IIE Trans. Healthcare Systems Engrg.* 5(2):101–123.

Svirsko AC, Norman BA, Rausch D, Woodring J (2019) Using mathematical modeling to improve the emergency department nurse-scheduling process. *J. Emergency Nursing* 45(4):425–432.

Twigg DE, Kutzer Y, Jacob E, Seaman K (2019) A quantitative systematic review of the association between nurse skill mix and nursing-sensitive patient outcomes in the acute care setting. *J. Adv. Nursing* 75(12):3404–3423.

Van Hulst D, den Hertog D, Nuijten W (2017) Robust shift generation in workforce planning. *Comput. Management Sci.* 14:115–134.

Wickert TI, Smet P, Berghe GV (2019) The nurse rerostering problem: Strategies for reconstructing disrupted schedules. *Comput. Oper. Res.* 104:319–337.

Verification Letter

Ajay Kumar, MD, FACP, SFHM, MBA, Executive Vice President and Chief Clinical Officer, Hartford HealthCare, writes:

"We are delighted to submit a verification letter for the paper titled "Optimization Automates Nurse Scheduling at Hartford Hospital Emergency Department" to the *INFORMS Journal on Applied Analytics*.

"The authors of the paper comprise of teams from Hartford HealthCare (HHC), academia, and a consulting firm, who have collaborated on the project since 2020. The study employs a combination of optimization, software development, and decision support methodologies to automate nurse scheduling at the emergency department (ED).

"The innovative solution has been successfully implemented in the daily operations of Hartford Hospital (HH), one of the largest teaching hospitals in New England (867 beds). HH is the flagship facility of HHC, the largest hospital network in Connecticut. Since deployed in March 2023, the practical software has become a key decision-making tool currently utilized by 200+ ED nurses and managers at HH.

"This successful implementation has brought substantial benefits to HHC. The nurse scheduling tool provides a labor-free process from input collection to schedule output, improving patient coverage and nurse satisfaction with reduced cost. The automation streamlines a labor-intensive nurse scheduling process, which previously took more than 88 hours of manual work every 6-week staffing cycle. Altogether, we project significant financial benefits at HH.

"We thank you very much for your time and for your consideration."

Liangyuan Na completed her PhD in operations research at Massachusetts Institute of Technology in 2023. Her research interests lie at the intersection of machine learning, optimization, and healthcare operations analytics. Her work has been recognized by awards including the Best Paper Award at the IEEE International Conference on Digital Health and Honorable Mention in INFORMS Doing Good with Good OR.

Jean Pauphilet is an assistant professor of management science and operations at London Business School. His research centers around discrete and robust optimization, machine learning, and their applications to healthcare and sustainability. His work has been recognized by multiple awards, including the 2020 INFORMS George E. Nicholson and the 2021 INFORMS Pierskalla awards. He completed his PhD in operations research at Massachusetts Institute of Technology in 2020.

Ali Haddad-Sisakht is a data scientist at Dynamic Ideas LLC. He works on decision-making problems in areas such as healthcare and transportation. He obtained his PhD in industry engineering from Iowa State University.

Louis Raison is a data scientist at Dynamic Ideas LLC. He builds and deploys optimization algorithms in school, transportation, and healthcare sections. He obtained his MS degree in business analytics from Massachusetts Institute of Technology.

Audrey Silver is a senior clinical operations manager of the emergency department at Hartford Hospital. She obtained her MSN in nursing management and executive leadership and holds NE-BC and CEN certifications.

Patricia Veronneau is a nurse director of emergency department and director of acute patient and family transitions at Hartford Hospital.

Nicole Vogt is a nurse manager of the emergency department at Hartford Hospital. She possesses a BS in nursing, along with multiple certifications in this field.

Dimitris Bertsimas is the associate dean of business analytics, Boeing professor of operations research, and faculty director of the Master of Business Analytics program at the Massachusetts Institute of Technology (MIT). His research focuses on optimization and machine learning, with practical applications. He is the recipient of numerous awards, including the von Neumann Theory Prize and INFORMS President Award. He completed his PhD in operations research and applied mathematics at MIT.