Title
A Conceptual Framework for Improving Critical Care Patient Flow and Bed Use

Journal
Annals of the American Thoracic Society, 12(6)

ISSN
2329-6933

Authors
Mathews, Kusum S
Long, Elisa F

Publication Date
2015-06-01

DOI
10.1513/annalsats.201409-419oc

Peer reviewed

A Conceptual Framework for Improving Critical Care Patient Flow and Bed Use

Kusum S. Mathews[1,2] and Elisa F. Long[3]

[1]Division of Pulmonary, Critical Care, and Sleep Medicine, Department of Medicine, and [2]Department of Emergency Medicine, Icahn School of Medicine at Mount Sinai, New York, New York; and [3]UCLA Anderson School of Management, Los Angeles, California

ORCID ID: 0000-0002-8810-2794 (K.S.M.).

Abstract

Rationale: High demand for intensive care unit (ICU) services and limited bed availability have prompted hospitals to address capacity planning challenges. Simulation modeling can examine ICU bed assignment policies, accounting for patient acuity, to reduce ICU admission delays.

Objectives: To provide a framework for data-driven modeling of ICU patient flow, identify key measurable outcomes, and present illustrative analysis demonstrating the impact of various bed allocation scenarios on outcomes.

Methods: A description of key inputs for constructing a queuing model was outlined, and an illustrative simulation model was developed to reflect current triage protocol within the medical ICU and step-down unit (SDU) at a single tertiary-care hospital. Patient acuity, arrival rate, and unit length of stay, consisting of a "service time" and "time to transfer," were estimated from 12 months of retrospective data (n = 2,710 adult patients) for 36 ICU and 15 SDU staffed beds. Patient priority was based on acuity and whether the patient originated in the emergency department. The model simulated the following hypothetical scenarios: (*1*) varied ICU/SDU sizes, (*2*) reserved ICU beds as a triage strategy, (*3*) lower targets for time to transfer out of the ICU, and (*4*) ICU expansion by up to four beds. Outcomes included ICU admission wait times and unit occupancy.

Measurements and Main Results: With current bed allocation, simulated wait time averaged 1.13 (SD, 1.39) hours. Reallocating all SDU beds as ICU decreased overall wait times by 7.2% to 1.06 (SD, 1.39) hours and increased bed occupancy from 80 to 84%. Reserving the last available bed for acute patients reduced wait times for acute patients from 0.84 (SD, 1.12) to 0.31 (SD, 0.30) hours, but tripled subacute patients' wait times from 1.39 (SD, 1.81) to 4.27 (SD, 5.44) hours. Setting transfer times to wards for all ICU/SDU patients to 1 hour decreased wait times for incoming ICU patients, comparable to building one to two additional ICU beds.

Conclusions: Hospital queuing and simulation modeling with empiric data inputs can evaluate how changes in ICU bed assignment could impact unit occupancy levels and patient wait times. Trade-offs associated with dedicating resources for acute patients versus expanding capacity for all patients can be examined.

Keywords: intensive care unit; resource allocation; queuing theory; computer simulation

(Received in original form September 9, 2014; accepted in final form March 4, 2015)

Supported by NHLBI award number 1K12HL109005-01 (K.S.M.).

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Heart, Lung, and Blood Institute or the National Institutes of Health.

Author Contributions: K.S.M. and E.F.L. both made substantial contributions to the conception or design of the work or the acquisition, analysis, or interpretation of data for the work; drafted and revised the work critically for important intellectual content; and had final approval of the version to be published. Both are in agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Correspondence and requests for reprints should be addressed to Kusum S. Mathews, M.D., M.P.H., Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, Box 1232, New York, NY 10029. E-mail: kusum.mathews@mssm.edu

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org

Ann Am Thorac Soc Vol 12, No 6, pp 886–894, Jun 2015

DOI: 10.1513/AnnalsATS.201409-419OC
Internet address: www.atsjournals.org

Optimizing intensive care unit (ICU) use is fundamental to quality improvement efforts by critical care physicians. In certain hospitals, particularly large academic centers, demand for ICU beds may outstrip supply, evidenced by a 32% increase in emergency department (ED) length of stay for critically ill patients between 2001 and 2009 (1), despite increases in the number of ICU beds nationwide (2). Reduced ICU bed availability can adversely affect hospital-wide patient throughput, especially within the ED and postsurgical care areas, and increase mortality of critically ill patients due to prolonged wait times for ICU bed assignment (3, 4). Strategies to address capacity strain include unit expansion, revision of triage policies, and targeted efforts to reduce throughput delays (5–7). A step-down unit (SDU), or intermediate care-level unit, can relieve ICU congestion by caring for lower-acuity patients in alternative settings (8–11).

Balancing the opposing goals of minimizing admission wait time for critically ill patients and maximizing bed use should be tailored to individual hospitals' priorities, patient population, and physical and financial constraints. Projected ICU use suggests that bed reconfiguration could help alleviate bed shortages (12). However, physically reconfiguring beds and observing the resulting impact on outcomes is time consuming, costly, and potentially catastrophic to patients, and therefore unwise to implement without clear evidence of its perceived benefit. Alternatively, a tailored computer simulation model can easily examine how different bed allocation or triage scenarios impact patient-centered outcomes. Simulation modeling and queuing theory are well-established methodologies used to improve hospital capacity planning (12–16). Prior studies have developed models to simulate varying hospital unit sizes, nurse or physician staffing levels, and different triage, discharge, and bed assignment policies to examine the impact on bed use, wait times, lengths of stay, readmissions, and mortality (17–22). Although many models capture patient flow through multidisciplinary ICUs (23), often with elective admissions (24), no studies included both an ICU and SDU with patients with different acuity levels within a single model.

This study's aim is to present a framework for understanding and improving patient flow through the medical ICU and SDU, and to present illustrative analysis for a single hospital. We describe the model's key assumptions, data analysis, and proposed outcomes. Our study offers insights for critical care physicians and hospital leaders aiming to better allocate limited ICU bed resources more effectively. Some of the results of the study have been previously reported in the form of abstracts (25, 26).

Methods

Before examining alternative bed allocation or triage strategies, a model of existing patient flow through a hospital unit, such as an ED or ICU, should be created. In general, queuing models must specify the arrival process, duration of service, and number of servers (e.g., beds, available staff); additional model components can be tailored for each specific setting. The main data requirements are patient-level throughput data and census or occupancy data for the specific hospital unit.

Within our ICU and SDU context, a queuing model includes five essential inputs: (*1*) number of beds for each unit, (*2*) patient type and priority level, (*3*) timing of patient arrival, (*4*) patient prioritization for admission, and (*5*) unit length of stay (Figure 1). To illustrate these elements and demonstrate queuing models' analytic value, we present data and simulation results from a single institution's medical ICU and SDU. Definitions and assumptions are given in Table 1. Full details on the methods used in model development can be found in the online supplement. Patient-level characteristics and hospital operations data (including locations and timestamps for admission and transfers) were collected for all patients admitted to the ICU and SDU (both of which primarily treat Medicine service patients, aged 18 years and older) over a 12-month period (June 2010 to June 2011). Data were collected via an automated query from Sunrise Clinical Manager and AllScripts (formerly Eclipsys) Sunrise Patient Flow and Bed Management database, with entry data for 20% of visits verified against written logs and chart documentation of Medical ICU/SDU admissions and discharges.

The institution's Human Investigations Committee approved this medical record review study before initiation. All data were gathered retrospectively.

Number of Beds

We first identify the number of beds for the unit(s) under consideration. As many ICUs use SDUs to augment capacity, both units should be represented. Our illustrative analysis was for a tertiary, academic, and community hospital with a 51-bed ward composed of 36 medical ICU beds and 15 medical SDU beds. These two units received admissions and transfers from the hospital's ED and inpatient Medicine wards, with a small fraction admitted from outside hospitals.

Patient Types

Admitted patients can be described in different ways, based on service (e.g., medicine, surgery), acuity, location of origin, diagnosis group, etc. In our setting, medicine service patients were categorized into one of four priority classes based on the observed triage policies (1 = acute non-ED, 2 = acute ED, 3 = subacute non-ED, and 4 = subacute ED), where 1 is the highest priority and 4 is lowest. This is a retrospective classification, which was not used clinically to triage patients at this institution but rather to model patient flow for this illustrative analysis. Acuity class was determined through detailed chart review (*see* Table E1 in the online supplement). Patients were assigned "subacute" status if they met the following criteria: (*1*) direct SDU admission, or (*2*) ICU admission for less than 24 hours, without a critical care–level diagnosis or receipt of a critical care–level intervention (27, 28). All other patients were designated as "acute" in severity, including those who expired within 24 hours of ICU admission. Although simplistic, this grouping allowed for straightforward comparisons when alternative strategies were later modeled.

Timing of Patient Arrivals

A patient "arrival" refers to the point when a physician requests ICU/SDU admission, before actual "service" (e.g., care within the unit) commences. The overall arrival rate can be estimated from patient-level timestamp data and is calculated as the reciprocal of the average time between consecutive patient arrivals. For each patient, ICU/SDU admission "wait time" is defined as the duration of time a patient spent in the ED or medical wards after a physician's request for ICU/SDU admission, before physically entering the unit (Figure 1).

For our illustrative model, all patient admissions or transfers into the ICU or SDU were considered separate patient visits; readmissions to the hospital or ICU (i.e., "bounce-backs") were included in the overall arrival rate. Patients treated in both the ICU and SDU appear as two separate observations in our dataset. ICU patients

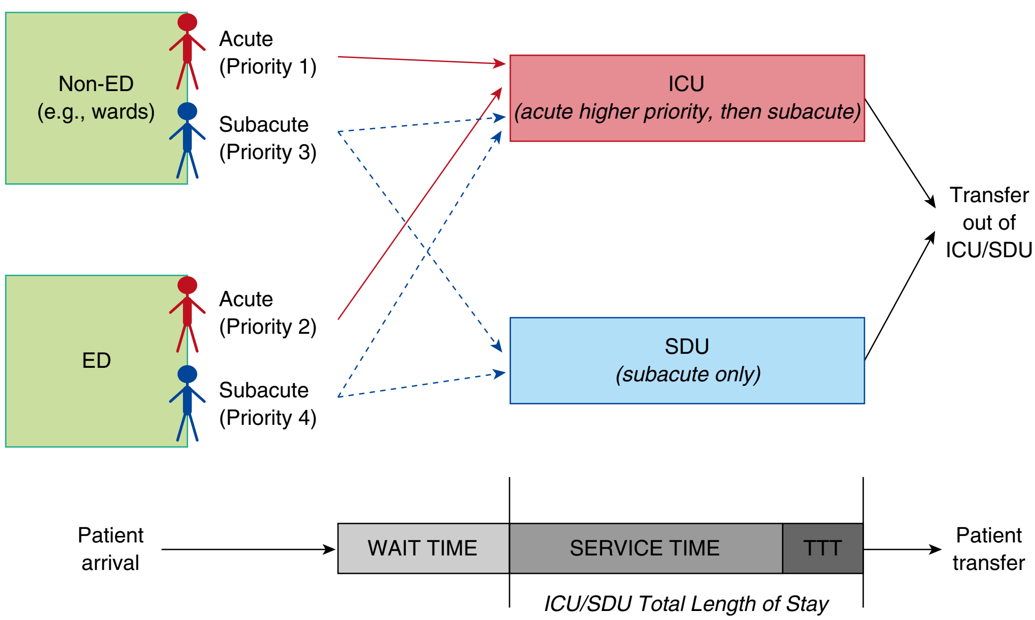


Figure 1. This overview of the priority queuing model displays the two different bed types (intensive care unit [ICU], accepting all patient classes, and step-down unit [SDU], accepting only subacute classes 3–4) and the flow of patients entering and exiting the units. Patients of differing acuity classes (1–4) originate in either the emergency department [ED] or non-ED locations (e.g., medical wards). These patients are shown from arrival, through their ICU/SDU length of stay—which is composed of service time (interval from entrance to unit until request for transfer out of unit) and time to transfer (TTT; interval from request for transfer out of unit to physical exit from unit)—and then actual patient transfer.

who "step down" to the SDU are modeled as exiting the ICU and immediately rearriving to the SDU.

ICU Triage Policy

The most complicated model component is the accurate representation of the ICU triage algorithm. Significant variation exists between institutions regarding admission criteria (29). Some hospitals consider acuity, diagnosis, and/or likelihood to benefit, whereas others use a "first-come, first-served" strategy.

The modeled medical ICU used a triage protocol incorporating the Society for Critical Care Medicine Guidelines for Admission, Discharge, and Triage for determining the acceptance and prompt transfer of critically ill patients (27). As part of hospital-wide throughput improvement efforts, minimizing ED length of stay was prioritized (30), prompting admission of intermediate-risk ED patients to the ICU if no SDU beds were available. Medicine ward patients with clinical deterioration requiring SDU or ICU transfer were prioritized over similar ED patients, as respiratory and nursing support were limited on the wards. Reflecting the institution's current protocol, the simulation model permitted ICU beds to be used by acute or subacute patients, and SDU beds were used by subacute patients only.

Length of Stay

Although many studies measure ICU length of stay as a single value, in many hospitals, length of stay can consist of two discrete periods (31) (*see* Figure 1). On physical transfer to an ICU/SDU bed, "service time" commences and continues until a request to transfer out of the unit. The service time distribution within our model was fit to past data and differed based on priority class. Due to wide heterogeneity in ICU service times, we further classified patients as "long-stay" or "short-stay" to better fit empirical distributions. Acute patients were (*1*) short-stay, if service was less than 2 weeks; or (*2*) long-stay, if longer than 2 weeks, clinically reflecting patients requiring more extensive critical care resources (e.g., prolonged mechanical ventilation, chronic critical illness, etc.). Subacute patients were (*1*) short-stay, if service time was less than 24 hours; or (*2*) long-stay, if greater than 24 hours.

Immediately after service time ends, a second period denoted as "time to transfer" (TTT) begins while the patient remains in the ICU or SDU awaiting transfer to a medical ward bed. During this period, care by medical providers continues, although the patient has been deemed clinically ready for a lower level of care. Based on clinical experience, we hypothesized that TTT might be affected by census levels within the ICU/SDU and on the wards. We used multivariate regression analysis to examine this relationship and modeled TTT as the same for ICU-ward transfers and SDU-ward transfers (Table E2). The simulation model generated each patient's TTT based on concurrent census levels, to capture the impact of ICU/SDU congestion on length of stay by accelerating TTT, as opposed to using fixed values from retrospective data.

Model Implementation

By sampling from the best-fitting empirical distributions, the simulation model randomly generated a set of variables for each hypothetical patient, including their priority class, arrival date and time, service time, and TTT. The model then assigned a particular bed based on availability and priority class. If no bed was immediately available, the patient queued until a bed was vacant. In our model, this process was performed for 2,000 patients, or approximately 6 months of ICU/SDU throughput, with calculation of several aggregate performance metrics.

Table 1. Key definitions and model assumptions for a medical intensive care unit/step-down unit

Term	Definition	Assumptions
Bed types		
ICU bed	ICU bed that can treat all acuity classes	Beds are identical and always staffed
SDU bed	SDU bed that can treat lower acuity patients	Beds are identical and always staffed
Patient characteristics		
Acute	Those with a critical care–level diagnosis (27) or who require a critical care–level intervention (28)	Patients originating from non-ED locations are higher priority
Subacute	Those admitted directly to SDU or admitted to the ICU who do not either have a critical care–level diagnosis or receive a critical car-–level intervention	Patients originating from non-ED locations are higher priority
Short-stay	Those who stay less than 2 wk (acute patients) or 24 h (subacute patients)	Classification used to estimate service time distribution from past data
Long-stay	Those who stay longer than 2 wk (acute patients) or 24 h (subacute patients)	Classification used to estimate service time distribution from past data
Model inputs		
Arrival rate	Number of patients per hour who require an ICU or SDU bed	Estimated from past data on patient throughput; calculated as the reciprocal of the average time between consecutive patient arrivals
Service time	Duration of time spent receiving care in the ICU or SDU	Estimated by fitting probability distributions to past data; can vary by acuity level and short or long stay
Time to transfer	Duration of time between request for transfer out of ICU or SDU and physical transfer to wards	Estimated using multivariate regression; can depend on concurrent patient census levels in the ICU/SDU and medical wards
Model outcomes		
Wait time	Duration of time between ICU or SDU bed request and physical transfer out of unit	Shorter wait times are preferred; can be validated against past data
Bed occupancy	Number of bed-hours that are occupied divided by total available bed-hours	Higher occupancy is more profitable but results in lower bed availability

Definition of abbreviations: ICU = intensive care unit; SDU = step-down unit.

The model implemented a priority queue, which assigned beds on a first-come, first-served basis, allowing for higher-acuity patients to "bump" those ahead in the queue. Once patients began their unit stay, a nonpreemptive policy was in effect (i.e., later patients could not displace existing patients in service). The bed assignment algorithm is detailed in the online supplement.

The model was validated by comparing model-simulated wait times for admission to those obtained from the retrospective cohort, by priority class, to ensure that the model closely represented patient flow through the ICU/SDU.

Simulated Bed Allocation Scenarios

In addition to simulating patient flow under the existing bed assignment policy, the model can be adjusted to explore different bed allocation strategies, which we illustrate with four broad policy or operational changes. First, we considered alternative partitions of the existing 51 staffed beds into separate ICU and SDU, ranging from 31 ICU + 20 SDU beds, up to 51 ICU + 0 SDU beds. Next, we considered a triage scenario holding constant the total bed number (36 ICU + 15 SDU) but with a cut-off policy that reserved the last unoccupied ICU bed for acute patients only. This was repeated for a scenario of two reserved beds. Third, unit expansion scenarios were simulated, where the 51-bed unit was enlarged by up to 4 additional ICU beds, with the SDU capacity remaining fixed at 15 beds. Finally, we simulated a hypothetical scenario (with 36 ICU + 15 SDU beds) where TTT was no longer census dependent but instead varied from a constant value of 6 hours down to 1 hour, to illustrate how improvements in transfer times could improve patient throughput.

Outcomes

Hospitals may wish to target various performance metrics as the outcome of interest. We identify two potential outcomes of consideration. Wait time for ICU/SDU admission by priority class is clinically relevant, as critically ill patients may deteriorate if care is delayed (32, 33). Second, average ICU/SDU bed occupancy levels—the proportion of time that beds are occupied—directly impacts profitability (34, 35). For each bed allocation scenario, we calculated each outcome across 10,000 simulation iterations.

Results

Patient Characteristics

Between June 2010 and June 2011, 3,165 patients were admitted to the ICU and 873 to the SDU. Among the ICU population, 14.4% were identified as subacute and alternatively could have been cared for in the SDU. Full patient characteristics and admission wait times were obtained for 2,710 patients. Service time, TTT, and corresponding census data were obtained for 1,694 patients (Table 2).

Table 2. General patient characteristics, grouped by acuity class

Total	Acute (ICU only) (*N* = 2,099)	Subacute (SDU + ICU Subacute) (*N* = 611)
ED origin, n (%)	1166 (55.6)	186 (30.4)
Age, mean ± SD, yr	61.36 ± 17.93	61.86 ± 18.13
Men, n (%)	1,078 (51.4)	303 (49.6)
Insurance, n (%)		
Medicare	1,176 (56.0)	340 (55.6)
Medicaid	464 (22.1)	143 (23.4)
Private	428 (20.4)	118 (19.3)
Self-pay/other	31 (1.5)	10 (1.6)
Admission wait times, median (IQR), h		
	1.05 (0.63–1.78)	1.20 (0.75–2.14)
ED	1.23 (0.85–1.85)	1.23 (0.78–1.85)
Non-ED	0.79 (0.47–1.63)	1.18 (0.71–2.29)
ICU/SDU length of stay, median (IQR), h*		
Total	51.73 (29.64–100.11)	24.93 (17.38–47.67)
Service time	39.73 (22.32–84.42)	15.92 (11.22–15.92)
Time to transfer	5.87 (2.98–14.01)	5.69 (3.23–11.62)
Hospital length of stay, median (IQR), d	8.01 (4.42–16.92)	6.20 (3.24–13.71)
In-hospital mortality, n (%)	329 (15.7)	56 (9.2)
Hospital readmission within 30 d, n (%)	471 (22.4)	137 (22.4)

Definition of abbreviations: ED = emergency department; ICU = intensive care unit; IQR = interquartile range; SDU = step-down unit.
*ICU/SDU length of stay was captured for 1,694 patients (62.5%), 1,479 (87.3%) of whom were classified as acute and 215 (12.7%) as subacute.

The mean time between consecutive patient arrivals across all priority classes was 1.96 hours, resulting in an average arrival rate of 0.51 patients per hour, or approximately 12 patients per day (Figure E1). Approximately 58% of arrivals occurred between 7 A.M. and 7 P.M. Among acute patients, 98% stayed less than 2 weeks, with a mean service time of 52.3 hours (Table 3). A lognormal distribution best fit service time data for these patients. We similarly empirically fit distributions for other patient classes (Figure E2).

Simulation Results: Existing Bed Allocation

Simulating the current structure of 36 ICU + 15 SDU beds resulted in a mean wait time for all patients of 1.13 hours (acute, 0.84 h; subacute, 1.39 h). Service time averaged 79.45 hours, and TTT averaged 3.73 hours. The simulated admission wait times by priority class were broadly similar to median values obtained from the retrospective cohort (Figure 2). Simulated service time and TTT values closely matched historical data (Figure E3).

Simulation Results: Alternative Policies

Reallocating SDU beds for ICU use—holding constant the total number of beds at 51—showed modest improvement in overall patient wait time as ICU size increased. With 31 ICU + 20 SDU beds, wait time averaged 1.23 hours; with 51 ICU beds, wait time decreased by 14% to 1.06 hours. However, the effect on individual priority classes differed substantially (Figure 3). Acute patients (priority 1–2) experienced a 52% reduction in wait time over this range in ICU bed allocation, whereas subacute patients (priority 3–4) suffered a 23% increase in wait time as fewer dedicated SDU beds were available. Detailed model results for each scenario are shown in Table 4.

As more SDU beds were reallocated as ICU beds, average occupancy rates in both units increased. For example, ICU bed use increased from 80% (31 ICU + 20 SDU beds) to 84% (51 ICU beds), suggesting that ICU patients spent more time in a bed and less time waiting for admission.

Reserving the last available ICU bed for acute patients (assuming 36 ICU + 15 SDU beds) reduced wait times for these patients to 0.20 hours (priority 1) and 0.42 hours (priority 2), a 63% improvement. In contrast, wait times for subacute patients worsened to 4.27 hours (priority 3 and 4). With two reserved beds, wait times improved by a further 11% to 0.14 hours (priority 1) and 0.29 hours (priority 2), whereas wait times for subacute patients worsened by 40% (5.98 h for priorities 3 and 4).

Table 3. Service time distribution based on empiric data

	Acute Patients	Subacute Patients
Short-stay duration	≤2 wk	≤24 h
Proportion, %	98	27
Mean (SD), h	50.9 (53.3)	12.3 (6.5)
Median (IQR), h	33.7 (19.6–63.5)	13.3 (9.3–18.0)
Fit distribution	Lognormal	Uniform
Long-stay duration	>2 wk	>24 h
Proportion, %	2	73
Mean (SD), h	645.4 (387.3)	131.6 (112.7)
Median (IQR), h	499.3 (398.0–760.0)	90.2 (60.0–180.6)
Fit distribution	Exponential	Exponential

Definition of abbreviation: IQR = interquartile range.

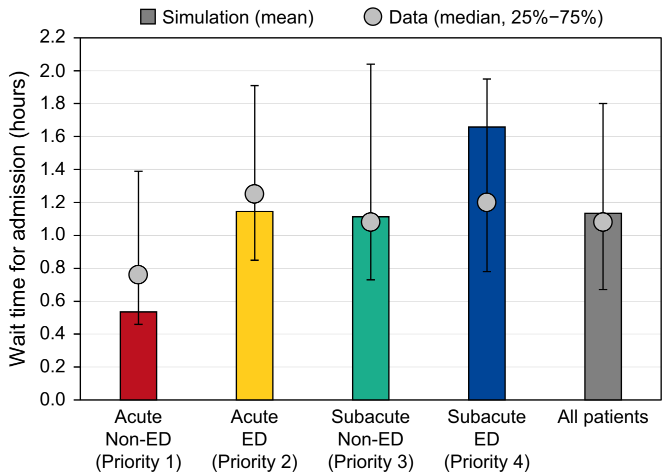


Figure 2. Mean wait times generated for each priority class within the simulation are represented as the colored bars. The historical data's median wait times with associated interquartile ranges are displayed with the *circle* and *error lines*. (Median values selected for comparison due to highly skewed nature of the retrospective wait time distributions.) Overall, the model's simulation output is similar to the median values seen in the empiric data, suggesting adequate calibration. The model generated longer average wait times for subacute emergency department (ED) patients (priority class 4) than seen in the empiric data, which may reflect a preference for bed assignment for patients with longer wait times.

The third simulation scenario considered a reduction in TTT for all patients and demonstrated linear improvements in bed occupancy rates and wait times. Reducing TTT from current levels to a target time of 1 hour for every patient led to a 46% improvement in wait times for bed assignment. This increased process throughput also freed up beds more quickly, reflected by the decrease in mean ICU bed occupancy from 80 to 66% with a 1-hour TTT target.

Finally, ICU bed expansion scenarios from one up to four extra beds substantially improved average wait times from 1.13 hours (status quo) to 0.76 hours (one additional bed) or 0.21 hours (four additional beds) for all patients. With this expansion, ICU bed occupancy modestly decreased from 79 to 76%, but SDU bed occupancy minimally changed (0.3% improvement).

Discussion

By presenting a conceptual framework and illustrative analysis for simulating patient flow through the ICU and SDU, this study demonstrated how queuing theory with data-driven inputs offers insights into improving ICU bed availability and admission wait time. Our empirically derived queuing model simulated ICU and SDU patient admissions and transfers, reliably replicated empiric wait times, and examined the effect of distinct bed allocation strategies and flexible triage policies on wait time and occupancy, tailored for a specific hospital.

The usefulness of queuing theory, a well-described analytical tool in operations management, has been documented in healthcare settings with practical applications for physicians and hospital

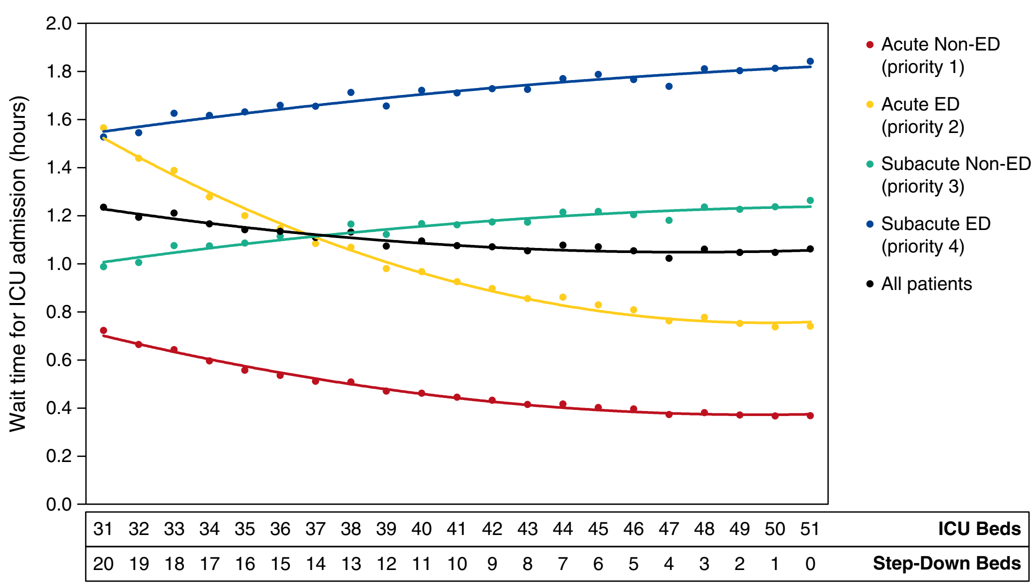


Figure 3. Average wait times for admission to the intensive care unit (ICU) or step-down unit (SDU) are shown for each scenario of ICU-SDU size (varying from 31 to 51 ICU beds, and 20 down to zero SDU beds). Wait times are grouped by priority class 1–4 and also in aggregate (*black line*). As the simulation was adjusted to have an increasing number of ICU beds reallocated from the SDU, wait times for all patients decreased on average but increased for subacute patients (priority class 3 and 4). ED = emergency department.

Table 4. Results of simulated bed allocation scenarios

Bed Scenario	Wait Time (*h*)			Bed Occupancy (%)		
	Overall	Acute	Subacute	Overall	ICU	SDU
Unit composition						
51 ICU + 0 SDU	1.06 (1.04–1.09)	0.56 (0.54–0.58)	1.55 (1.52–1.59)	84.3 (84.3–84.4)	84.3 (84.3–84.4)	—
46 ICU + 5 SDU	1.05 (1.03–1.08)	0.60 (0.58–0.62)	1.48 (1.45–1.52)	84.3 (83.3–84.4)	83.1 (83.0–83.1)	95.7 (95.7–95.8)
41 ICU + 10 SDU	1.07 (1.05–1.10)	0.68 (0.67–0.70)	1.44 (1.40–1.47)	84.3 (84.2–84.3)	81.7 (81.7–81.8)	94.7 (94.7–94.7)
36 ICU + 15 SDU*	1.13 (1.11–1.16)	0.84 (0.82–0.86)	1.39 (1.35–1.42)	84.2 (84.1–84.2)	80.4 (80.4–80.5)	93.2 (93.2–93.2)
31 ICU + 20 SDU	1.23 (1.20–1.26)	1.14 (1.12–1.17)	1.26 (1.22–1.29)	84.0 (83.9–84.0)	79.6 (79.5–79.7)	90.7 (90.7–90.8)
Reserved acute beds						
1 Reserved bed	2.23 (2.17–2.28)	0.31 (0.31–0.32)	4.27 (4.16–4.38)	84.7 (84.7–84.8)	80.8 (80.8–80.9)	94.0 (94.0–94.1)
2 Reserved beds	3.00 (2.93–3.07)	0.22 (0.21–0.22)	5.99 (5.85–6.12)	84.8 (84.7–84.8)	80.8 (80.7–80.8)	94.4 (94.4–94.4)
ICU bed expansion						
37 ICU + 15 SDU	0.76 (0.74–0.78)	0.58 (0.56–0.59)	0.92 (0.89–0.94)	83.0 (82.9–83.0)	78.9 (78.8–79.0)	93.0 (93.0–93.0)
38 ICU + 15 SDU	0.50 (0.48–0.51)	0.39 (0.37–0.40)	0.59 (0.57–0.61)	82.0 (81.9–82.0)	77.7 (77.6–77.7)	92.9 (92.9–92.9)
39 ICU + 15 SDU	0.33 (0.32–0.34)	0.26 (0.25–0.27)	0.38 (0.37–0.40)	81.1 (81.1–81.2)	76.6 (76.6–76.7)	92.8 (92.8–92.8)
40 ICU + 15 SDU	0.21 (0.20–0.22)	0.17 (0.16–0.18)	0.24 (0.23–0.25)	80.5 (80.4–80.5)	75.9 (75.8–76.0)	92.7 (92.7–92.7)
Target TTT						
1 h	0.69 (0.67–0.70)	0.51 (0.50–0.53)	0.83 (0.81–0.85)	73.5 (73.3–73.6)	65.5 (65.3–65.7)	92.7 (92.6–92.7)
2 h	0.85 (0.83–0.87)	0.64 (0.62–0.66)	1.03 (1.00–1.06)	74.7 (74.5–74.8)	67.1 (66.9–67.3)	92.9 (92.8–92.9)
3 h	1.06 (1.04–1.09)	0.79 (0.77–0.81)	1.30 (1.26–1.33)	75.5 (75.4–75.7)	68.2 (68.0–68.4)	93.1 (93.1–93.1)
4 h	1.28 (1.25–1.30)	0.94 (0.92–0.97)	1.56 (1.52–1.60)	76.6 (76.4–76.7)	69.6 (69.4–69.8)	93.3 (93.2–93.3)
5 h	1.57 (1.54–1.60)	1.16 (1.13–1.18)	1.93 (1.88–1.97)	77.6 (77.4–77.7)	70.9 (70.7–71.1)	93.5 (93.5–93.5)
6 h	1.96 (1.91–2.00)	1.44 (1.40–1.48)	2.40 (2.35–2.46)	78.6 (78.4–78.8)	72.3 (72.1–72.5)	93.7 (93.7–93.7)

Definition of abbreviations: ICU = intensive care unit; SDU = step-down unit; TTT = time to transfer.
Data presented as mean (95% confidence interval). Mean values are reported based on a simulation with 10,000 iterations; 95% confidence intervals are calculated assuming a "sample size" of 10,000. Wait time is time from ICU bed request to physical transfer. Bed occupancy is proportion of available bed-hours that are occupied by patients.
*Existing number of ICU and SDU beds at this institution.

leadership to better understand patient throughput and improve bed availability (36–38). Our model was informed by the current admission process at one institution but could be readily adapted to other locations with adjustments of empiric inputs and bed assignment algorithm. Similar to past ICU models maximizing potential health benefits or reserving unit beds for elective surgeries (14, 39, 40), we examined bed allocation strategies and flexible triage policies that can alleviate shortages during periods of high ICU demand. This model adds to the literature by incorporating both ICU and SDU bed types, using a priority bed assignment algorithm that paralleled actual clinical decision making and providing an excellent testing ground for system and practice changes.

The simulation results demonstrate some fundamental insights from queuing theory—which could apply across institutions—as well as some specific conclusions for our institution. Pooling beds, or eliminating different bed types, is a well-established queuing paradigm to reduce waiting, although of course effect size depends on the specific parameters of the hospital setting. In our setting, pooling the ICU and SDU into one unit may be preferred, when expansion of both units is not feasible due to physical or economic constraints. Reallocating SDU beds into one ICU unit (all 51 beds) improves bed use while reducing wait times, because it maximizes flexibility in permitting all beds to serve all patient types. In the original allocation (36 ICU + 15 SDU), it is possible for an acute patient to wait for ICU admission during high census, while an SDU bed is simultaneously unoccupied. When we eliminate SDU beds, pooling facilitates faster admission of acute patients.

There are potential shortcomings to pooling ICU/SDU beds, however. For instance, with a single pooled unit, staffing levels would likely change as the nursing-to-patient ratio is higher in an ICU than in an SDU, potentially increasing hospital costs. On the other hand, under the current system, subacute patients in the ICU use more nursing and specialist resources than clinically necessary, which is also costly (41–43). Before implementing such bed reconfigurations, a hospital could conduct a careful cost–benefit analysis to better measure these trade-offs. Flexible staffing models that better match patient type with needed resources (e.g., nursing and respiratory support) may generate further cost savings (44).

Our analyses highlighted another opportunity for improvement: reduction of TTT, an ineffective use of resources where patients occupy ICU beds while waiting for transfer to medical wards (31). We observed substantial improvements in admission wait times and overall bed availability with faster transfer times out of the ICU and SDU, regardless of unit sizing. For example, reducing TTT from current levels to 1 hour—which would involve centralized quality improvement effects affecting all patient flow from ED presentation to hospital discharge (30, 45)—generated improved wait times equivalent to building one to two additional ICU beds, a costly solution that would only temporarily relieve congestion without added patient benefit (46, 47). Moreover, reduced TTT was the only simulated scenario that achieved ICU bed occupancy levels below 70%, which is an increasingly important metric as the demand for ICU beds continues to grow. Nevertheless, reducing TTT would likely require operational changes, including additional transfer staff, greater coordination between the ICU and wards, and improved timeliness of hospital discharges, without increasing the risk of delays for ED patients awaiting admission to wards or early discharges for non-ICU patients (17, 18, 30).

Given the established link between delays in ICU admission and increased risk of mortality (3, 4), admission wait time is a clinically supported performance metric for our model (48). Although many hospitals focus on decreasing wait times for admission from the ED, critical care physicians and hospital administrators can evaluate additional performance metrics, such as ICU/SDU occupancy rates, before concluding that ICU expansion is needed (49, 50).

Strengths of our modeling approach are its use of empirical data over a 12-month period for two types of units (ICU and SDU) and an adaptive bed assignment algorithm that considers priority bed assignment to accommodate more critically ill patients earlier. The model also accounts for the "speed up" of TTT occurring during periods of higher ICU census, reflecting real-world observations of expedited patient transfer times when available ICU beds are limited.

Limitations of our model include the absence of a more granular severity of illness estimation in retrospective data. Although the acute/subacute and long/short-stay categorizations were based on chart documentation and published guidelines, there may be additional factors affecting acuity classification and admission priority. Although we aggregated subacute patients, these lower-risk patients admitted to the ICU may reflect a different clinical class than those admitted to the SDU (10, 51). We also had limited sample sizes for some patient subgroups, causing our distribution fitting for short-stay subacute patients to appear less precise than for other subgroups. Due to limited data on patient timestamps resulting from completely missing weeks of hospital operations reports, as well as only 9 months of ICU and medicine ward census data, our original sample size reflecting all admissions to the ICU/SDU over 1 year is reduced. However, we do not observe any systematic bias in the missing data. Historical census data were available for only four time points per day, which prevents us from examining how time-series trends in floor crowdedness affect ICU TTT. Future models could incorporate alternative staffing models, more granular measures of patient severity, and other triage policies similar to the bed reservation strategy in times of surge.

Using real-world patient-level data for a major academic U.S. hospital, we provided an illustrative framework to use queuing theory and simulation modeling to understand and improve ICU and SDU patient flow. We examined hypothetical unit resizing scenarios and measured the impact on admission wait times and bed occupancy rates. Before implementing large-scale bed reallocation plans or triage policies, such as switching SDU beds to ICU permanently or temporarily in times of ICU surge, such policies can be examined within simulation models such as ours, with adjustments made for local institutional policies and characteristics, to measure the resulting benefits and tradeoffs to patients and the health system as a whole. Small changes in bed assignment policies can have a significant effect on operational metrics, with potentially profound effects on patient health outcomes. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

References

1 Herring AA, Ginde AA, Fahimi J, Alter HJ, Maselli JH, Espinola JA, Sullivan AF, Camargo CA Jr. Increasing critical care admissions from U.S. emergency departments, 2001-2009. *Crit Care Med* 2013;41: 1197–1204.

2 Wallace DJ, Angus DC, Seymour CW, Barnato AE, Kahn JM. Critical care bed growth in the United States: a comparison of regional and national trends. *Am J Respir Crit Care Med* 2015;191:410–416.

3 Cardoso LT, Grion CM, Matsuo T, Anami EH, Kauss IA, Seko L, Bonametti AM. Impact of delayed admission to intensive care units on mortality of critically ill patients: a cohort study. *Crit Care* 2011;15:R28.

4 Chalfin DB, Trzeciak S, Likourezos A, Baumann BM, Dellinger RP; DELAY-ED study group. Impact of delayed transfer of critically ill patients from the emergency department to the intensive care unit. *Crit Care Med* 2007;35:1477–1483.

5 McConnell KJ, Richards CF, Daya M, Bernell SL, Weathers CC, Lowe RA. Effect of increased ICU capacity on emergency department length of stay and ambulance diversion. *Ann Emerg Med* 2005;45:471–478.

6 Howell E, Bessman E, Marshall R, Wright S. Hospitalist bed management effecting throughput from the emergency department to the intensive care unit. *J Crit Care* 2010;25:184–189.

7 Nguyen YL, Wunsch H, Angus DC. Critical care: the impact of organization and management on outcomes. *Curr Opin Crit Care* 2010;16:487–492.

8 Eachempati SR, Hydo LJ, Barie PS. The effect of an intermediate care unit on the demographics and outcomes of a surgical intensive care unit population. *Arch Surg* 2004;139:315–319.

9 Thungjaroenkul P, Kunaviktikul W. Possibilities for cost containment in intensive care. *Nurs Health Sci* 2006;8:237–240.

10 Prin M, Wunsch H. The role of stepdown beds in hospital care. *Am J Respir Crit Care Med* 2014;190:1210–1216.

11 Nasraway SA, Cohen IL, Dennis RC, Howenstein MA, Nikas DK, Warren J, Wedel SK; American College of Critical Care Medicine of the Society of Critical Care Medicine. Guidelines on admission and discharge for adult intermediate care units. *Crit Care Med* 1998;26:607–610.

12 Green L. Capacity planning and management in hospitals. *Operations Research and Health Care* 2005;70:15–41.

13 Cochran JK, Roche K. A queuing-based decision support methodology to estimate hospital inpatient bed demand. *J Oper Res Soc* 2008;59: 1471–1482.

14 Yang M, Fry MJ, Raikhelkar J, Chin C, Anyanwu A, Brand J, Scurlock C. A model to create an efficient and equitable admission policy for patients arriving to the cardiothoracic ICU. *Crit Care Med* 2013;41:414–422.

15 Zhu Z, Hen BH, Teow KL. Estimating ICU bed capacity using discrete event simulation. *Int J Health Care Qual Assur* 2012;25:134–144.

16 McManus ML, Long MC, Cooper A, Litvak E. Queuing theory accurately models the need for critical care resources. *Anesthesiology* 2004; 100:1271–1276.

17 Powell ES, Khare RK, Venkatesh AK, Van Roo BD, Adams JG, Reinhardt G. The relationship between inpatient discharge timing and emergency department boarding. *J Emerg Med* 2012;42: 186–196.

18 Diwas Singh KC, Terwiesch C. An econometric analysis of patient flows in the cardiac intensive care unit. *Manufacturing and Service Operations Management* 2012;14:50–65.

19 Dobson G, Lee HH, Pinker E. A model of ICU bumping. *Oper Res* 2010; 58:1564–1576.

20 Chan CW, Farias VF, Bambos N, Escobar GJ. Optimizing intensive care unit discharge decisions with patient readmissions. *Oper Res* 2012; 60:1323–1341.

21 Li X, Beullens P, Jones D, Tamiz M. An integrated queuing and multi-objective bed allocation model with application to a hospital in China. *J Oper Res Soc* 2009;60:330–338.

22 Khare RK, Powell ES, Reinhardt G, Lucenti M. Adding more beds to the emergency department or reducing admitted patient boarding times: which has a more significant influence on emergency department congestion? *Ann Emerg Med* 2009;53:575–585.
23 Barado J, Guergué JM, Esparza L, Azcárate C, Mallor F, Ochoa S. A mathematical model for simulating daily bed occupancy in an intensive care unit. *Crit Care Med* 2012;40:1098–1104.
24 Kolker A. Process modeling of ICU patient flow: effect of daily load leveling of elective surgeries on ICU diversion. *J Med Syst* 2009;33:27–40.
25 Mathews KS, Long EF. Impact of intensive care unit re-sizing on admission wait times: a priority queuing simulation model for acute and sub-acute patients [abstract]. *Am J Respir Crit Care Med* 2014; 189:A3726.
26 Mathews KS, Jenq GY, Pisani MA, Long EF. Characterizing the flow of short-stay patients in the intensive care unit. Presented at the Society for Medical Decision Making 34th Annual Conference. October 17-20, 2012, Phoenix, AZ.
27 Task Force of the American College of Critical Care Medicine, Society of Critical Care Medicine. Guidelines for intensive care unit admission, discharge, and triage. *Crit Care Med* 1999;27:633–638.
28 Keene AR, Cullen DJ. Therapeutic Intervention Scoring System: update 1983. *Crit Care Med* 1983;11:1–3.
29 Walter KL, Siegler M, Hall JB. How decisions are made to admit patients to medical intensive care units (MICUs): a survey of MICU directors at academic medical centers across the United States. *Crit Care Med* 2008;36:414–420.
30 Jweinat J, Damore P, Morris V, D'Aquila R, Bacon S, Balcezak TJ. The safe patient flow initiative: a collaborative quality improvement journey at Yale-New Haven Hospital. *Jt Comm J Qual Patient Saf* 2013;39:447–459.
31 Howell MD. Managing ICU throughput and understanding ICU census. *Curr Opin Crit Care* 2011;17:626–633.
32 Gorunescu F, McClean SI, Millard PH. Using a queueing model to help plan bed allocation in a department of geriatric medicine. *Health Care Manage Sci* 2002;5:307–312.
33 Hagen MS, Jopling JK, Buchman TG, Lee EK. Priority queuing models for hospital intensive care units and impacts to severe case patients. *AMIA Annu Symp Proc* 2013;2013:841–850.
34 Tierney LT, Conroy KM. Optimal occupancy in the ICU: a literature review. *Aust Crit Care* 2014;27:77–84.
35 Costa AX, Ridley SA, Shahani AK, Harper PR, De Senna V, Nielsen MS. Mathematical modelling and simulation for planning critical care capacity. *Anaesthesia* 2003;58:320–327.
36 Troy PM, Rosenberg L. Using simulation to determine the need for ICU beds for surgery patients. *Surgery* 2009;146:608–617. [Discussion, pp. 617–620.]
37 Kim SH, Chan CW, Olivares M, Escobar GJ. ICU admission control: an empirical study of capacity allocation and its implication on patient outcomes. Columbia Business School Research Paper. 2014;12:34.
38 Alavi-Moghaddam M, Forouzanfar R, Alamdari S, Shahrami A, Kariman H, Amini A, Pourbabaee S, Shirvani A. Application of queuing analytic theory to decrease waiting times in Emergency Department: does it make sense? *Arch Trauma Res* 2012;1: 101–107.
39 Shmueli A, Sprung CL, Kaplan EH. Optimizing admissions to an intensive care unit. *Health Care Manage Sci* 2003;6:131–136.
40 Kim SC, Horowitz I, Young KK, Buckley TA. Flexible bed allocation and performance in the intensive care unit. *Journal of Operations Management* 2000;18:427–443.
41 Edbrooke DL, Minelli C, Mills GH, Iapichino G, Pezzi A, Corbella D, Jacobs P, Lippert A, Wiis J, Pesenti A, *et al*. Implications of ICU triage decisions on patient mortality: a cost-effectiveness analysis. *Crit Care* 2011;15:R56.
42 Delgado MK, Liu V, Pines JM, Kipnis P, Gardner MN, Escobar GJ. Risk factors for unplanned transfer to intensive care within 24 hours of admission from the emergency department in an integrated healthcare system. *J Hosp Med* 2013;8:13–19.
43 Christian MD, Fowler R, Muller MP, Gomersall C, Sprung CL, Hupert N, Fisman D, Tillyard A, Zygun D, Marshal JC; PREEDICCT Study Group. Critical care resource allocation: trying to PREEDICCT outcomes without a crystal ball. *Crit Care* 2013;17:107.
44 Iapichino G, Radrizzani D, Rossi C, Pezzi A, Anghileri A, Boffelli S, Giardino M, Mistraletti G, Bertolini G; GiViTI Group. Proposal of a flexible structural-organizing model for the Intensive Care Units. *Minerva Anestesiol* 2007;73:501–506.
45 Chaboyer W, Lin F, Foster M, Retallick L, Panuwatwanich K, Richards B. Redesigning the ICU nursing discharge process: a quality improvement study. *Worldviews Evid Based Nurs* 2012;9:40–48.
46 Wunsch H. Is there a Starling curve for intensive care? *Chest* 2012;141: 1393–1399.
47 Wunsch H, Wagner J, Herlim M, Chong DH, Kramer AA, Halpern SD. ICU occupancy and mechanical ventilator use in the United States. *Crit Care Med* 2013;41:2712–2719.
48 Gorunescu F, McClean SI, Millard PH. A queueing model for bed-occupancy management and planning of hospitals. *J Oper Res Soc* 2002;53:19–24.
49 Terwiesch C, Diwas KC, Kahn JM. Working with capacity limitations: operations management in critical care. *Crit Care* 2011;15:308.
50 Wild C, Narath M. Evaluating and planning ICUs: methods and approaches to differentiate between need and demand. *Health Policy* 2005;71:289–301.
51 Junker C, Zimmerman JE, Alzola C, Draper EA, Wagner DP. A multicenter description of intermediate-care patients: comparison with ICU low-risk monitor patients. *Chest* 2002;121: 1253–1261.

Online Data Supplement

A Conceptual Framework for Improving Critical Care Patient Flow and Bed Utilization

Kusum S. Mathews, MD MPH
kusum.mathews@mssm.edu

Elisa F. Long, PhD
elisa.long@anderson.ucla.edu

We developed a novel priority queuing simulation model to mimic the flow of patient through a Medical Intensive Care Unit (ICU) and Step-Down Unit (SDU), and applied it to a single tertiary-care hospital with an existing 51-bed unit, consisting of 36 ICU beds and 15 SDU beds. Patients arrive to the ICU/SDU from the Emergency Department (ED) or non-ED locations.

This document outlines our key assumptions and methodology, which consists of six components: patient variables, bed assignment algorithm, empirical analysis of patient-flow data, model implementation, validation, and simulated bed allocation scenarios.

A. Patient variables

The model is a discrete-event simulation model, which simulates the flow of a hypothetical cohort of 2,000 patients through an ICU and SDU. The "events" for each patient include arrival to the ICU/SDU, initiation of service, transfer request to the Medical wards, and physical unit discharge. By fitting probability distributions to past data, we are able to randomly generate a set of variables for each patient. The simulation then randomly draws from these probability distributions, so each simulated patient obtains a unique set of timestamps that are consistent with historical data.

The following variables are simulated for each patient *i*:

Variable	Possible values	Description	Reference
Interarrival time	X_i ~ *Exponential*	Each patient incurs a randomly generated interarrival time (the reciprocal of the arrival rate), which is used to compute the exact arrival date and time of consecutive patients.	Figure E1
Priority class	$P_i \in \{1, 2, 3, 4\}$	Each patient is randomly assigned one of four priority classes based on historical proportions. The class refers to whether the patient is acute/subacute and originating in the ED/non-ED.	Table E1
Duration of stay	$D_i \in$ {*short*, *long*}	Each patient is randomly assigned to be short- or long-stay based on historical proportions in each acuity level.	Table 3
Service time	Y_i ~ *Lognormal*, *Exponential*, or *Uniform*	Refers to the duration of active service in the ICU/SDU, and depends on acute/subacute status and short/long stay.	Figure E2
Time-to-transfer	$Z_i \in R^+$	At the time of service completion for each patient, the Medicine floor census is randomly drawn from nine months of historical data. This refers to the number of floor beds that are occupied, and is used to compute time-to-transfer for each patient.	Table E2

Note: R^+ refers to the set of positive real numbers.

B. Bed assignment algorithm

Priority queue and bed assignment

Upon arrival to the ICU (i.e., when the physician requests a bed), patients are assigned a bed on a first-come, first-served basis within priority class. A higher priority patient (e.g., priority 1) will be assigned a bed before a lower priority patient (e.g., priority 4), even if the lower priority patient arrived first. If a bed is available when a patient arrives – and no higher priority patients are waiting – then the patient is immediately assigned a bed. If both ICU and SDU beds are available at the time of arrival, subacute patients (priority 3-4) will preferentially be assigned an SDU bed to maintain ICU bed availability, whereas acute patients (priority 1-2) must always be assigned an ICU bed because of their more intensive resource needs. If no beds are available, then the patient must wait in the queue and be assigned according to their priority class.

Later patients arriving to the ICU could "bump" earlier patients in queue (before service commences) if (a) the later patient was strictly higher priority, (b) the next available bed was adequate, and (c) the earlier patient had been bumped fewer than $k=10$ times. For example, a class 4 (subacute) patient could be bumped by classes 1-3 if an ICU bed became available, and only by class 3 if a SDU bed became available. The model's constraints of priority classes and algorithm for "bumping" algorithm were designed to mimic current clinical ICU admitting practices at the hospital, with the k limit established so that subacute patients could not be indefinitely waiting if no beds became available. We chose the k patient limit rather than a specified time limit because the priority model considers each new patient arrival an "event", and we could thus count the number of later arriving patients who are eligible to bump an earlier patient in the queue. A limit of k=10 corresponds to approximately two nursing shifts after the initial patient arrival, and our clinical experience suggests that few patients, if any, queue for an ICU bed for longer than this.

The following section details the algorithm used to assign a bed for each patient *i*.

[1] Compute the arrival time (α_i) as the sum of the previous patient's arrival time (α_{i-1}) and the interarrival time (X_i):

$$\alpha_i = \alpha_{i-1} + X_i$$

[2] Determine the set of adequate beds (B_i) appropriate for the newly arriving patient:

$$B_i = \begin{cases} ICU & if\ P_i = 1,2 \\ ICU, SDU & if\ P_i = 3,4 \end{cases}$$

Identify the sub-set of beds that are unoccupied upon patient arrival:

a) If exactly one adequate bed is available at time α_i, assign that bed to patient *i*. Continue to Step 6.

b) If more than one bed type is available at time α_i, assign the lower tier bed:

$$Assign\ patient\ i \begin{cases} ICU & if\ P_i = 1,2 \\ SDU & if\ P_i = 3,4 \end{cases}$$

Continue to Step 6.

c) If no beds are available at time α_i, tentatively assign patient *i* the first-available bed, b_1, with a projected time of availability:

$$\tau_1 = \min\{\tau_j : j \in B_i\}$$

[3] Compute the arrival time (α_{i+1}) and set of capable beds (B_{i+1}) for the next arriving patient *i*+1.
If all of the following conditions hold, assign bed b_1 to patient *i*+1 instead of patient *i*:

a) Patient *i*+1 arrives before patient *i* is tentatively scheduled to begin service in bed b_1

$$\alpha_{i+1} < \tau_i$$

b) Bed b_1 is also suitable for patient *i*+1 acuity level

$$b_1 \in B_{i+1}$$

c) Patient *i*+1 is strictly higher priority than patient *i*

$$P_{i+1} < P_i$$

[4] If Step 3 conditions (a-c) do not hold for patient *i*+1, repeat for patient *i*+2, *i*+3, …*i*+10.

[5] If Step 3 conditions (a-c) do hold for some patient *h*, tentatively assign the second-available bed, b_2, for original patient *i* where:

$$\tau_2 = \min\{\tau_j : j \in B_i, j \neq b_1\}$$

Repeat Steps 3-4 using the second-available bed, third-available bed, etc., up to a limit of k bed bumps (k=10) for patient i.

Return to Step 1 for the next patient $i+1$.

Calculation of ICU/SDU length of stay

Once patients begin their unit stay, a non-preemptive policy is in effect, meaning that subsequent patients cannot displace an existing patient in a bed. This is the current policy at the hospital under consideration.

The following section details the algorithm used to determine ICU/SDU length of stay, which consists of a service time and time-to-transfer (TTT), for each patient. Based on our clinical observations, overall lengths of stay tend to be shorter during times of high ICU/SDU census; this is incorporated into the simulation through the census-dependent TTTs (described in Section C below).

[1] The bed (b) assigned to patient i, becomes available at time τ_b, which becomes the patient's service begin (β_i) time:

$$\beta_i = \tau_b$$

The patient's service time (Y_i) is randomly generated, and service end time (ξ_i) occurs when transfer to the Medicine ward is requested:

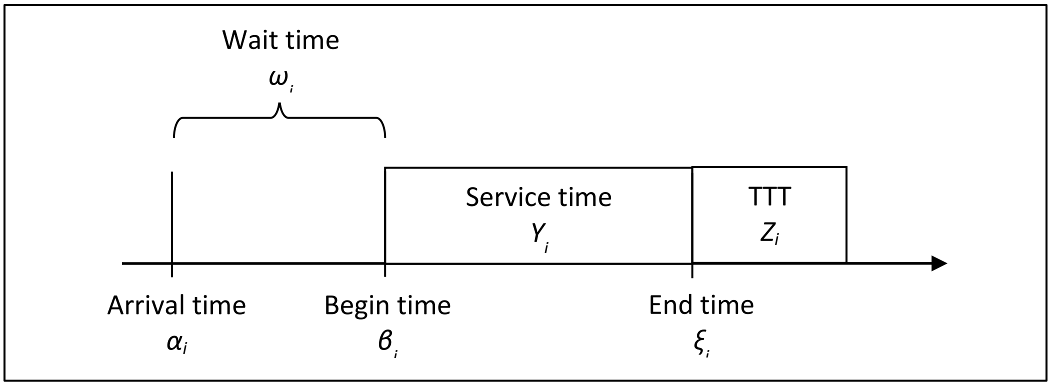
$$\xi_i = \beta_i + Y_i$$

The wait time (ω_i) for bed assignment for patient i is only positive if the service begin time occurs after the arrival time:

$$\omega_i = \max(\beta_i - \alpha_i, 0)$$

If the patient is immediately assigned a bed, then there is no wait time:

$$\alpha_i = \beta_i$$
$$\omega_i = 0$$



[2] Compute a time-to-transfer (Z_i) based on the ICU census level within the simulation model at time of transfer request (ξ_i) and the randomly sampled Medicine floor census, as well as other patient-level effects (Section C).

[3] After completion of TTT, patient *i* is transferred to the wards, and bed *b* becomes available to future patients.

C. Empirical analysis of patient-flow data

Priority class

In the simulation model, each patient's priority class is determined based on proportions observed over a 12-month period (June 2010 – June 2011). Acute patients are deemed more critically ill and are thus higher priority than subacute patients. Patients originating from non-ED areas are higher priority than ED patients because of the enhanced critical care resources available in the ED at our institution.

Table E1: Descriptions of patient classes and priorities

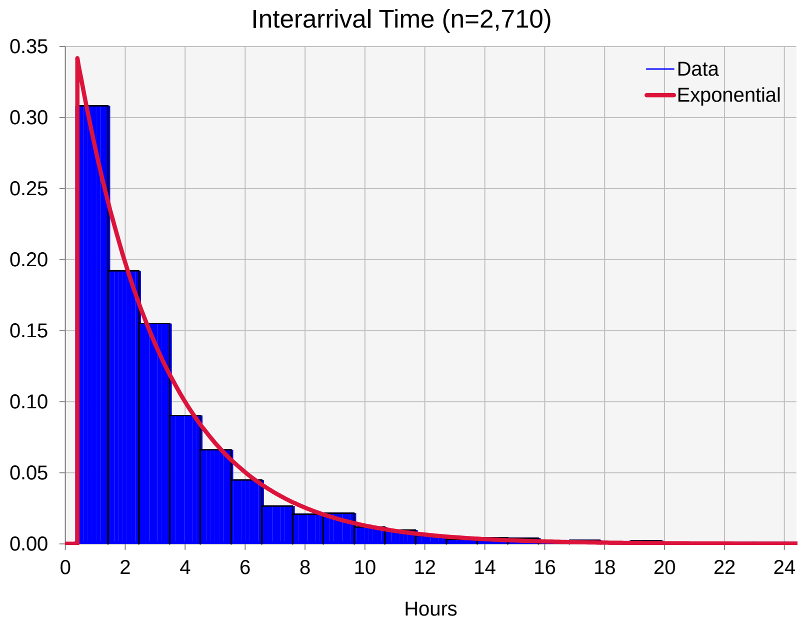
Priority	Acuity	Origin	Proportion
1	Acute*	Non-ED	23.7%
2		ED	28.2%
3	Subacute‡	Non-ED	21.0%
4		ED	27.1%

* ACUTE: Patients with ICU-level admission diagnosis(1) OR requiring ICU-level intervention(2)
‡ SUBACUTE: Patients meeting neither of the above criteria admitted to the ICU and/or patients admitted directly to the SDU

Interarrival time

The histogram below shows the interarrival times (intervals between consecutive ICU/SDU bed requests) of the historical cohort, with an overlying best-fitting exponential curve. As with most real-world queuing systems, the exponential distribution is the best fitting distribution in our setting. This implies that patient arrivals to the ICU/SDU follow a Poisson process. The mean interarrival time from this data is 1.96 hours (median = 1.36 hours). The arrival rate (0.51 patients/hour) is the reciprocal of the mean interarrival time (1.96 hours). This arrival rate amounts to approximately 12 patients admitted to the ICU/SDU daily. This overall arrival rate is then subdivided into the four priority classes based on the proportions given in Table E1.

Figure E1: Interarrival times for patients arriving to the ICU/SDU with best-fitting exponential curve

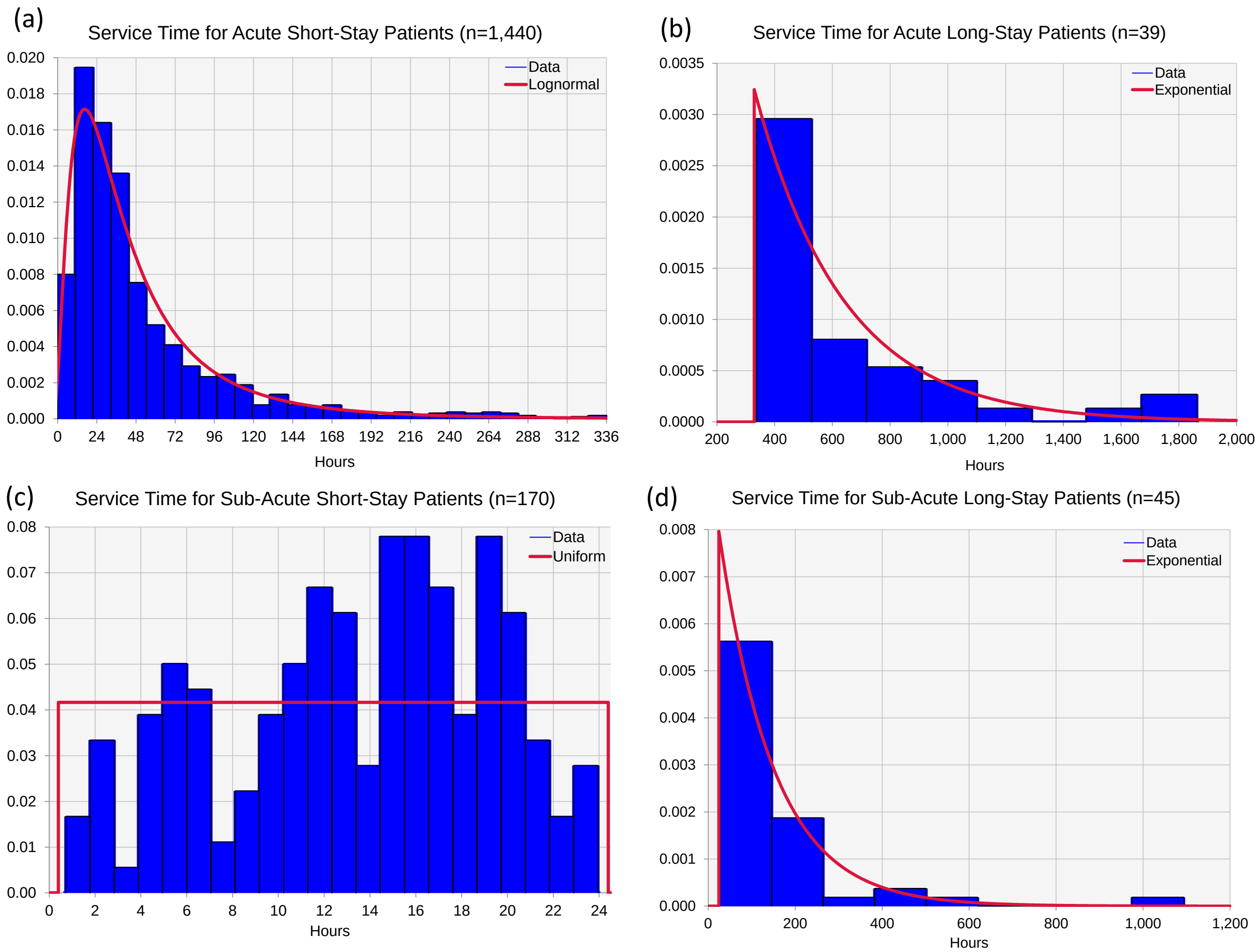


Service time

For each class of patients (acute short-stay, acute long-stay, subacute short-stay, subacute long-stay) we fit probability distributions to historical data on ICU/SDU service time (the period of active care before transfer to the Medicine wards is requested). The distinction of short-stay versus long-stay is based on clinical differences in these populations. We do not observe differences in service time for patients originating in the ED or non-ED.

We have the most complete patient flow data for acute short-stay patients (those who stayed less than two weeks), and we find that a Lognormal distribution best fits, based on the Akaike Information Criterion, a measure of statistical goodness-of-fit. The other distributions are similarly fit to past data, although we note that these are potentially less accurate given the paucity of data for these sub-populations.

Figure E2: Service times in the ICU/SDU with best-fitting probability distributions



(a) Acute short-stay patients (<14 days), all admitted to the ICU, with a lognormal best-fitting curve.
(b) Acute long-stay patients (>14 days), all admitted to the ICU, with exponential best-fitting curve.
(c) Subacute short-stay patients (<24 hours), admitted to SDU or ICU, with a uniform best-fitting curve.
(d) Subacute long-stay patients (>24 hours), admitted to SDU or ICU, with an exponential best-fitting curve.

Time-to-transfer

Prior studies have demonstrated that critical care may be accelerated during periods of strain, in order to free up beds for incoming patients.(3, 4) To examine whether service time or time-to-transfer (TTT) is potentially impacted by the unit's crowdedness, we utilized multivariate regression analysis using patient-level throughput data, along with census counts within the ICU/SDU and on the Medicine floors near the time a patient completed service. Our institution recorded census levels four times per day (midnight, 5am, 1pm, 9pm), both within the ICU/SDU and the Medicine wards.

We find that census levels have no significant effect on service time, although they do impact TTT, suggesting that patients' bed transfer times depend on the surrounding units' crowdedness. However, we find no effect of patients' location of origin (ED or non-ED) on duration of TTT. We present some of the key regression results in Table E2. To allow for nonlinear effects of ICU census (i.e., moving from 30 to 31 occupied ICU beds may have a different effect than from 50 to 51 beds, when the unit is completely full), we divided historical census levels into ten deciles. Within each decile, we examined whether service time (Model 1) or TTT (Model 2) differed, controlling for other patient factors, such as age, ED origin, or primary DRG code.

Because we are taking the natural logarithm of the dependent variables (service time or TTT), the coefficients represent percent-changes. The results suggest that ICU staff are speeding up the transfer of patients when the ICU is very full, to make room to accommodate newly arriving patients. Conversely, Medicine floor staff might be slowing the transfer of incoming ICU patients during periods of floor bed strain.

In order to generate a TTT for each hypothetical patient in the model, we compute ICU census levels within the model near the time of bed transfer request out of the ICU/SDU. For Medicine floor census levels, we randomly sample Medicine floor census levels from historical values. We population both census values into the regression equation, to generate a predicted time-to-transfer for the simulation model. Thus, the simulation model captures the impact of ICU and Medicine floor congestion on length of stay by accelerating the time-to-transfer.

For additional information on the regression results, please contact the authors directly.

Table E2: Regression model for ICU duration of stay (coefficients and standard errors)

		Model 1 Log(Service time)	Model 2 Log(Time-to-transfer)
ICUcensus	2ndDecile	−0.001 (0.117)	0.058 (0.122)
	3rdDecile	−0.055 (0.113)	0.092 (0.118)
	4thDecile	−0.148 (0.123)	0.185 (0.128)
	5thDecile	−0.151 (0.110)	−0.141 (0.115)
	6thDecile	0.085 (0.135)	0.242 (0.141)
	7thDecile	−0.080 (0.117)	−0.487 (0.122)***
	8thDecile	0.035 (0.117)	−0.390 (0.121)**
	9thDecile	−0.048 (0.124)	−0.562 (0.129)***
	10thDecile	−0.234 (0.139)	−0.990 (0.145)***
FLOORcensus	2ndDecile	−0.167 (0.121)	0.193 (0.126)
	3rdDecile	−0.088 (0.123)	0.221 (0.128)
	4thDecile	−0.040 (0.124)	0.459 (0.129)***
	5thDecile	−0.130 (0.127)	0.467 (0.132)***
	6thDecile	−0.086 (0.128)	0.667 (0.134)***
	7thDecile	−0.154 (0.130)	0.757 (0.136)***
	8thDecile	−0.065 (0.128)	0.931 (0.134)***
	9thDecile	−0.070 (0.134)	1.161 (0.139)***
	10thDecile	−0.127 (0.136)	0.949 (0.142)***
Medicaid		−0.152 (0.080)	−0.081 (0.084)
Medicare		−0.122 (0.075)	−0.086 (0.078)
Age		0.004 (0.002)	0.002 (0.002)
ED		−0.398 (0.057)***	−0.044 (0.060)
Weekday		−0.041 (0.063)	−0.003 (0.065)
Dayshift		0.285 (0.086)***	0.349 (0.090)***
Month		Included	Included
DRG		Included	Included
Observations		1,434	1,434
Adjusted R^2		0.16	0.19

Significance levels: 0.05 (*), 0.01 (**), 0.001 (***)
ICUcensus = Intensive Care Unit census; FLOORcensus = Medicine floor census; ED=Emergency Department; Weekday = Monday to Friday; Dayshift = 7am to 7pm; DRG = Diagnosis related group code
Included refers to independent categorical variables that were included in the regression, but the coefficients are not displayed due to space constraints.

D. Model implementation

The simulation model was implemented in Microsoft Excel, where each hypothetical patient was on a separate line in a spreadsheet. The relevant variables (interarrival time, service time, time-to-transfer) were generated for each patient, and the bed assignment algorithm was implemented by assigning the appropriate bed to each patient.

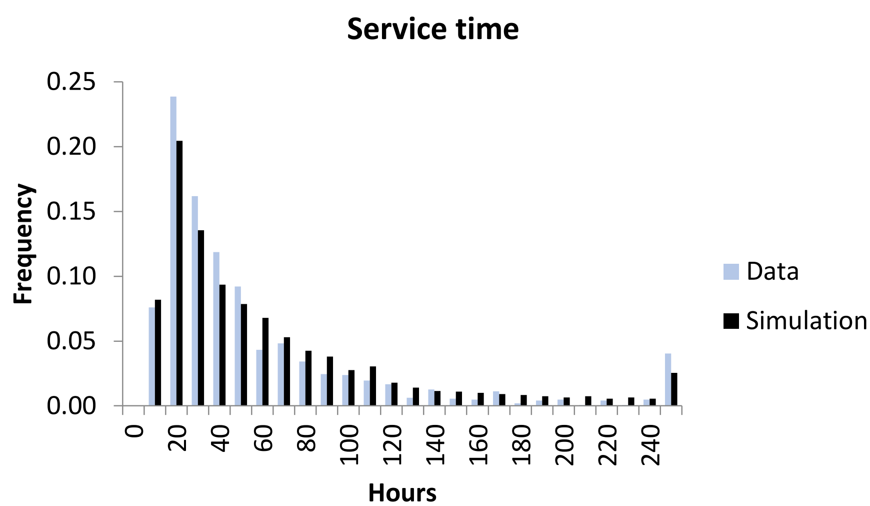
Probability distribution fitting was performed using @Risk software (Palisade Decision Tools). Multivariate regression analysis was performed in Stata.

E. Validation

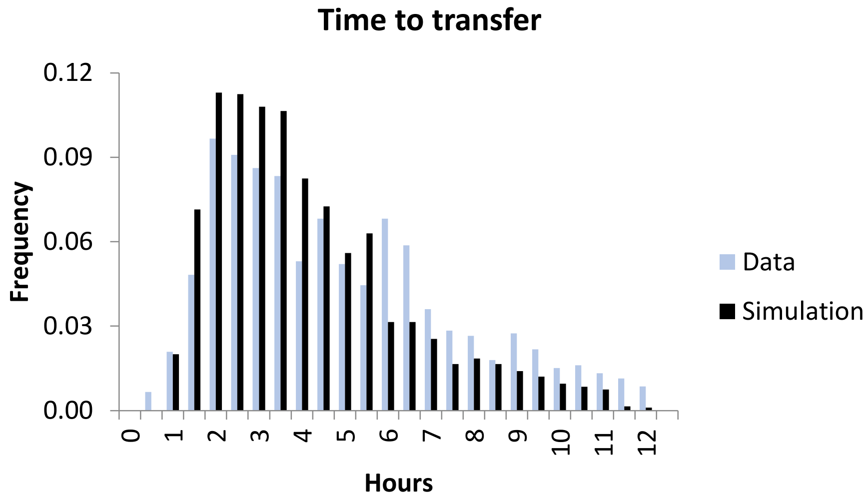
The model simulated patient flow through the ICU/SDU for 2,000 patients, representing about six months of throughput. We then repeated this process for 10,000 iterations, and we averaged outcome values over all iterations to obtain mean values. In Figure E4, we show a histogram of all service times and times to transfer based on historical data (blue bars), as well as one run of the simulation model (black bars). Note, each time the simulation is run, the histogram will change slightly due to random variation. We excluded extreme outliers (e.g., patients with transfer times exceeding 24 hours) because these are not clinically valid, and likely represent an error in timestamp coding.

Figure E3: Histograms of (a) service time and (b) time-to-transfer from 12-months of data and 1 run of queuing simulation model

(a)



(b)



In general, we find that the simulation model closely matches historical data, even with all of our assumptions regarding ICU and floor census levels, and priority bed assignment. Models tailored to other hospital settings should ensure that the model closely approximates the actual triage and bed assignment process at that specific institution.

F. Simulated bed allocation scenarios

After building and validating the simulation model, we first considered the status quo bed allocation at our institution (36 ICU + 15 SDU beds). We then considered "bed reallocation" of the existing 51 beds as follows:

- 51 ICU + 0 SDU beds
- 46 ICU + 5 SDU beds
- 41 ICU + 10 SDU beds
- 31 ICU + 20 SDU beds

We also simulated every scenario in between but we presented results for only these scenarios to highlight the relative impact on admission wait times and ICU/SDU bed occupancy.

Second, we simulated a "cut-off bed" policy, which maintained the current bed allocation (36 ICU + 15 SDU beds), but once 35 ICU beds were occupied and only 1 ICU bed remained open, we reserved it for only acute patients rather than either acute or subacute patients. This helped ensure that higher priority acute patients likely encountered an open bed upon arrival. This was also repeated using 2 reserved beds.

Third, we evaluated a "bed expansion" scenario, where we maintained the existing 15 SDU beds, but added between 1 and 4 additional ICU beds, which represents the current physical bed capacity at our institution.

Finally, we considered a “targeted time-to-transfer” scenario as a thought exercise. Here, we omitted the effects of ICU or Medicine ward census levels on TTT, but instead assigned a fixed value for all patients’ TTT, ranging from 1 to 6 hours. This was to explore whether such a goal could sufficiently reduce admission wait times, and we compared this to the far costlier policy of building additional ICU beds.

References

1. Guidelines for intensive care unit admission, discharge, and triage. Task Force of the American College of Critical Care Medicine, Society of Critical Care Medicine. *Crit Care Med* 1999;27:633-638.

2. Keene AR, Cullen DJ. Therapeutic Intervention Scoring System: update 1983. *Crit Care Med* 1983;11:1-3.

3. Kc DS, Terwiesch C. An econometric analysis of patient flows in the cardiac intensive care unit. *Manuf Serv Oper Manag* 2012;14:50-65.

4. Stelfox HT, Hemmelgarn BR, Bagshaw SM, Gao S, Doig CJ, Nijssen-Jordan C, Manns B. Intensive care unit bed availability and outcomes for hospitalized patients with sudden clinical deterioration. *Arch Intern Med* 2012;172:467-474.