Proceedings of the 2016 Winter Simulation Conference
T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds.

A SIMULATION MODEL OF PATIENT FLOW THROUGH THE EMERGENCY DEPARTMENT TO DETERMINE THE IMPACT OF A SHORT STAY UNIT ON HOSPITAL CONGESTION

John Chavis

Center for Applied Mathematics
Cornell University
206 Frank H.T. Rhodes Hall
Ithaca, NY 14853, USA

Amy L. Cochran

Department of Mathematics
University of Michigan
530 Church Street
Ann Arbor, MI 48109, USA

Keith E. Kocher

Department of Emergency Medicine
Medical School
University of Michigan
2800 Plymouth Rd.
NCRC Bldg 16
Ann Arbor, MI 48109, USA

Valerie N. Washington

Department of Systems and Industrial Engineering
Kennesaw State University
W. Clair Harris Textile Center
Room 108, MD 9061
1100 South Marietta Parkway
Marietta, GA 30060, USA

Gabriel Zayas-Cabán

Center for Healthcare Engineering & Patient Safety
University of Michigan
Industrial and Operations Engineering Building
1205 Beal Avenue
Ann Arbor, Michigan 48109, USA

ABSTRACT

One of the most critical and costly decisions made in emergency departments (EDs) is whether to admit a patient into the hospital. These decisions require investment in time for patient testing and treatment, delaying care to other patients. Short-stay units (SSUs) are an alternative to discharging or fully admitting ED patients, allowing extended patient observation. However, little is understood about the design of an SSU and its impact on outcomes and congestion. Here, we introduce a discrete-event simulation model of a hospital system (ED, inpatient units, and SSUs). By analyzing records from a tertiary teaching hospital, we determine realistic parameters and identify important features, such as triage level and processes depending on triage level, time, and congestion. We contend that performance metrics, e.g. time to first contact, critically depend on downstream hospital units. To demonstrate utility, we use the simulation model to assess bed occupancy over time.

1 BACKGROUND

Increasingly in the United States, most patients requiring unscheduled admission are hospitalized through the emergency department (ED). EDs are primarily designed for short-term acute care needs while inpatient hospital units are designed for acute needs requiring longer term care. Admitting patients to a hospital unit is one of the most expensive routine decisions made in healthcare (Sabbatini, Nallamothu, and Kocher 2014). Ideally, patients would be admitted or transferred only when needed and discharged as soon as they are stable enough to go home. However, this clinical assessment is not always immediately obvious and may be better revealed with additional time as the patient's medical condition progresses and results from further diagnostic testing are returned. EDs may respond to these situations by seeking a setting that provides a period of observation. Complicating the process, EDs commonly encounter episodes of crowding that can lead to long waits for care and adverse patient outcomes. Better informed admission and transfer decisions may improve outcomes and reduce costs without sacrificing timely access to care.

To address this problem, medical short-stay units (SSUs) have been proposed as a way to reduce crowding in EDs, while improving admission decisions (Centers for Medicare and Medicaid Services 2011, Damiani et al. 2011, Lovejoy and Desmond 2011). SSUs are a formal alternative to an admission to inpatient hospital units. They are designed to provide extended evaluation for up to 24 to 48 hours in order to determine if a patient is stable enough to be discharged home or requires admission to the inpatient hospital for additional treatment (Centers for Medicare and Medicaid Services 2011, Damiani et al. 2011, Lovejoy and Desmond 2011). SSUs provide an alternative disposition decision for ED patients who may benefit from further observation, such as those are not ill enough to be admitted, but not well enough to be discharged. However, whether SSUs improve outcomes, efficiency, or costs remains an open-ended and hotly-debated topic with important policy implications (Noel-Miller and Lind 2015, Zuckerman et al. 2016).

2 OBJECTIVE

To analyze the impact of SSUs on patient flow in the hospital, we introduce a robust discrete-event simulation of a stochastic service network motivated by the University of Michigan hospital system (UMHS). While our network model is based on our experience with UMHS, it includes many features common to hospitals elsewhere and can be tailored to a specific hospital system with ease. Certain features that distinguish our model from other hospital simulations (c.f. Saghafian et al. (2012), Saghafian et al. (2014), Shi et al. (2015)) include

1. The interaction between the ED and inpatient units such as SSUs.
2. Boarding: preparing patients or waiting for a transfer between units.
3. Overflow: transferring patients to a unit other than the one designated for their condition to reduce crowding.
4. Patient classes: designations used to determine decisions, processes, routing, among others.
5. Priority: prioritizing patients that are waiting for service based on their class.
6. Blocking: keeping a patient in a unit rather than transferring due to overcrowding.
7. Service times, arrival times, and transfers that depend on congestion, time, patient class, and the units.

To demonstrate the utility of the model, we use the simulation model to determine beds occupancy over time in the ED and in the SSU. Bed occupancy over time can be used to determine how to size SSUs by choosing the number of beds that achieves some target probability of delay (c.f. (Green 2006)). Clinical data from electronic healthcare records at the UMHS is analyzed to specify patient classes and realistic parameters in the simulation.

3 HOSPITAL PATIENT FLOW

Patients typically come to the ED without advance notice, and move into, through, and out of the ED according to needs and resource availability. Figure 1 (top) depicts an example of a patient journey at the

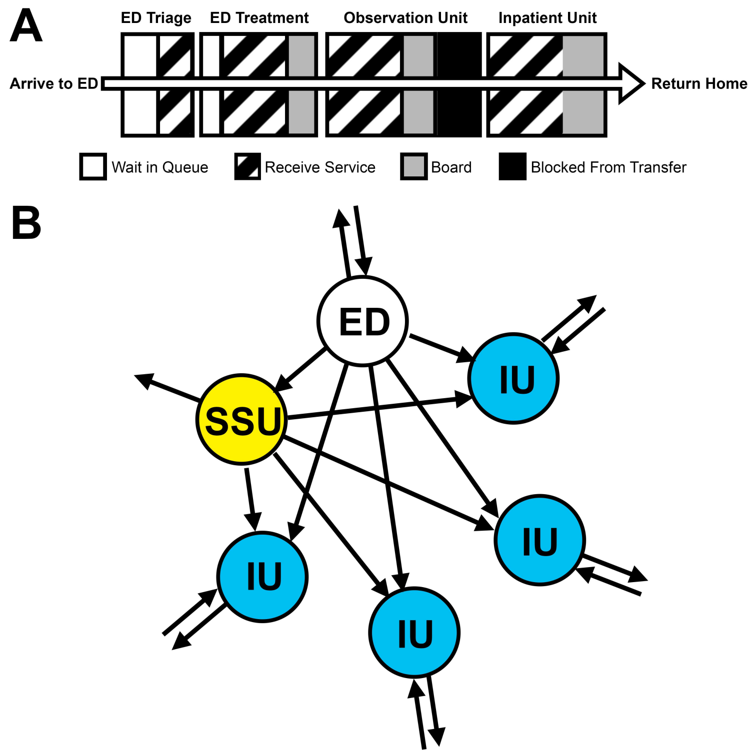

Figure 1: **(A)** An example of a patient journey arriving to the ED. As he/she travels to receive treatment between units and phases of service (ED triage and treatment, observation or SSU, and inpatient unit), a patient may also have to wait in a queue, board (i.e. prepare for a transfer), or be blocked from a transfer. **(B)** Typical routes patients can take within the UMHS. Each node represents a hospital unit, and arrows into and out of a node represent the flow of patients into and out of a unit. Patients can arrive to any hospital unit from outside the system, can move from the ED to the SSU, IU, or home, from the SSU to any IU or home, and from any IU to another IU or home.

UMHS starting from the ED. Although variations occur depending on the needs of the patients and the specific hospital under consideration, the patient flow starting in the ED can be described as follows.

The patient arrives to the ED, registers and immediately proceeds to triage, where she/he is assigned an acuity level by a triage nurse based on severity of illness. Acuity is typically assigned using the emergency severity index (ESI), a five-level triage algorithm. Higher acuity patients (1 or 2) are almost immediately brought to a bed for treatment. Lower acuity patients wait for treatment until either she/he leaves the ED without receiving treatment or is brought to a bed.

Once a lower-acuity patient is assigned to a room/bed within the department, they are visited by a provider. The providers then determine a plan of care which may involve diagnostic testing such as imaging (radiographs, ultrasound CT scans, MRI), laboratory work, and treatment. After the patient receives testing and treatment, he/she is either well enough to be discharged home or is admitted to the hospital.

During the admission process, the attending evaluates the patient status to decide whether the patient needs to be admitted into an inpatient unit (IU). If so, the patient information is reviewed by a case manager and information (e.g. estimated length of stay, diagnosis, special needs, precautions and level of care) is entered into the electronic healthcare records system. Once entered, and before the patient is officially accepted, there is a discussion and care hand off between the providers in the ED and physicians in the IU.

After the request is compiled, it is sent to the bed management staff who checks for bed availability. A bed in the IU is chosen according to an iterative process until a bed becomes available. ED staff inform the transportation staff to move the patient into the assigned IU beds, the patient is then transferred, and receives treatment in an IU.

Once in the IU, the patient is continuously assessed and treated until discharge, or, occasionally, transferred to another unit in the hospital. The patient's care team generally rounds on a daily basis and it is at this time when the decision to discharge the patient home or to transfer the patient to a different unit within the hospital is made. In the former case, discharge orders are sent and processed, and the patient leaves the hospital. Otherwise, providers at the receiving unit are notified of the transfer and the patient is moved to the new unit.

4 STOCHASTIC SERVICE NETWORK MODEL

Based on the hospital patient flow description from the previous section, we model the hospital system as a stochastic network. It consists of J nodes (hospital units), each with one or several servers (hospital beds). Patients arrive to any node from outside the network and belong to one of K classes. Patients arrive from outside the network according to a non-homogeneous Poisson process with rate that depends on the job class.

At each node, patients are processed in two phases. The first phase of service represents treatment up to the disposition decision. The second phase of service represents the time required for boarding when a patient is sent to another unit or final processing when a patient is sent home. The first phase of service begins when a patient reaches a server at the queue. At this time, a service time is generated randomly according to a probability distribution that depends on up to five factors: node, class, time, congestion, and where the job should have been routed (in the case of overflow). Here, congestion is the number of patients waiting for or in service.

Upon completion of the first phase of service, it is randomly determined whether the patient should be routed to another unit (admitted) or sent home (discharged). The routing is determined by a probability transition matrix which is class dependent. At the same time, an admission decision is made where to actually route the patient (allowing for bed managers to manage overflow). This decision can depend on node, class, time, and congestion.

After the first phase of service (once the disposition is made), another service time is generated randomly accordingly to a probability distribution that depends on the node, class, time, congestion, and where the job will actually be routed. For example, a patient that will be routed to the surgery unit can have a boarding time that depends on the surgery unit.

Upon completion of the second phase of service, three things could happen. First, the patient can be immediately routed home or to the next node (determined at disposition) provided there is a server available. Second, the patient can be blocked from service (i.e. remains in the same node), because there are no servers available. At which point, the patient is placed in a queue and must wait in their current location, keeping their current server occupied, until a server is made available for them. Third, the patient can be routed to the next node, but placed in a queue, until a server is made available for them. The placement in the queue, i.e. the job's priority, can depend on class, time, and congestion.

4.1 Model Justification

Our modeling assumptions are based on an empirical analysis of electronic healthcare records data from the University of Michigan Health System collected from August 2015 through October 2015 for the ED and from July 2015 in the IUs. This included around 17,799 patient encounters in the ED and on around 3,476 patient encounters in IUs. The following characteristics were available: clinical characteristics (e.g. triage level), and event dates and times at the ED, SSUs, and other IUs. This work is part of the study *Emergency Department Patient Flow: The Continuum of Care from Sick in the Emergency Department to Healthy in the Hospital to Home at the University of Michigan*, for which the University of Michigan's Biomedical Institutional Review Board (IRB) approved all research procedures (HUM00105968).

We distinguished six patient classes: five that arrive initially to the ED and one that arrives initially to the IUs. ED patient classes were based on acuity and disposition decision (Acuity level 1-2 Admit, Acuity 1-2 Discharge, Acuity 3 Admit, Acuity 3 Discharge, Acuity 4-5). The IU patient class consists of scheduled arrivals or transfers from another hospital. For each patient class, the arrival process was modeled as a non-homogeneous Poisson process with a piece-wise constant arrival rate function (Figure 2A). For each ED patient class, the arrival rate function was constant over 2-hour periods by day of the week. For the IU patient class, the arrival rate function was constant over 2-hour periods independent of day of the week (due to limited data; see bottom Figure 3A). The constants were estimated from the data as the average number patients in the respective class for the relevant 2-hour period. We tested our proposed arrival process model using statistical tests proposed by (Lewis 1965) based on the Kolmogorov-Smirnov (KS) statistic (Kim and Whitt 2014a, Kim and Whitt 2014b).

ED treatment time was defined as the time from when the patient is roomed in the ED to when the ED disposition is made. ED treatment time was analyzed by class, number of patients in or waiting for treatment, and time of day (Figures 2B–D). The likelihood function was used to find an appropriate one- or two-parameter family of continuous probability distribution for ED treatment times. This led to treatment times in the ED being modeled as lognormal random variables. In addition, average treatment time was found to depend on the patient class and ED occupancy (patients in service at the ED). Hence, parameters for the lognormal distribution were determined by maximum likelihood estimation, where for each ED class, the location parameter for the lognormal distribution was modeled as a linear function of ED congestion and the scale parameter was assumed to be constant.

ED boarding time was defined as the time from when the ED disposition is made to when the patient exits the ED. Using a similar analysis to ED treatment times led to boarding times in the ED being modeled as lognormal random variables that depend on where the patient was transferred (Figures 2E–F). Again, parameters were calculated for each IU using maximum likelihood estimation.

Five inpatients nodes or "units" were defined: Intensive Care Unit, Telemetry, General Care, Pediatric Unit, and Medical SSU. These definitions are used by the UMHS Admissions & Bed Coordination Center to divide physical units and services based on broad clinical resources available. Although not captured in detail in the model, it is important to note that across these nodes, there are approximately ten different inpatient services at the UMHS (Neuroscience, Head Neck and Plastics, Cardiovascular & Thoracic, Transplant (Med/Surg) & Surgical Services, Medical Services, Cancer, Short Stay, Moderate care, and ICU), which refers to the physician team caring for a patient, and approximately 60 physical IUs, which refers to locations in the hospital. The hospital also includes cardiovascular and other operating rooms, an ED overflow unit, a burn acute care unit, and a catheter laboratory. The operating rooms were excluded and considered an inpatient service. The other units were excluded because they had a very small number of beds.

Except for the SSU, IU treatment times were defined as the time from when the patient was admitted to the unit to when the patient was discharged. IU treatment times were modeled as lognormal random variables that depend on the unit (Figure 3B). For the SSU, treatment times were calculated as the difference between when the patient was admitted to the SSU and when the patient left the unit. SSU treatment times were also modeled as lognormal random variables (Figure 3B). Parameters were estimated using maximum likelihood estimation. Boarding times in an IU/SSU and boarding times for any patient sent home was assumed to be zero.

Patient routing was modeled as in the diagram in Figure 1. Patients can arrive to the ED or one of the IUs based on patient class (for the IU class, c.f. Figure 3E). From the ED, patients can be sent home or to the SSU or IUs (Figure 3C–D). From the SSU, patients can be sent home or sent to another IU (Figure 3F). From the IU (not the SSU), patients are sent home. For the ED classes, the transition probability matrix for routing out of the ED was estimated from the data as fractions of the total count of transitions. For the IU classes, the transition probability matrix for routing into an IU was estimated from the data as fractions of the total count, and similarly, the transition probability matrix for routing out of an SSU was also estimated from the data as fractions of the total count of transitions.

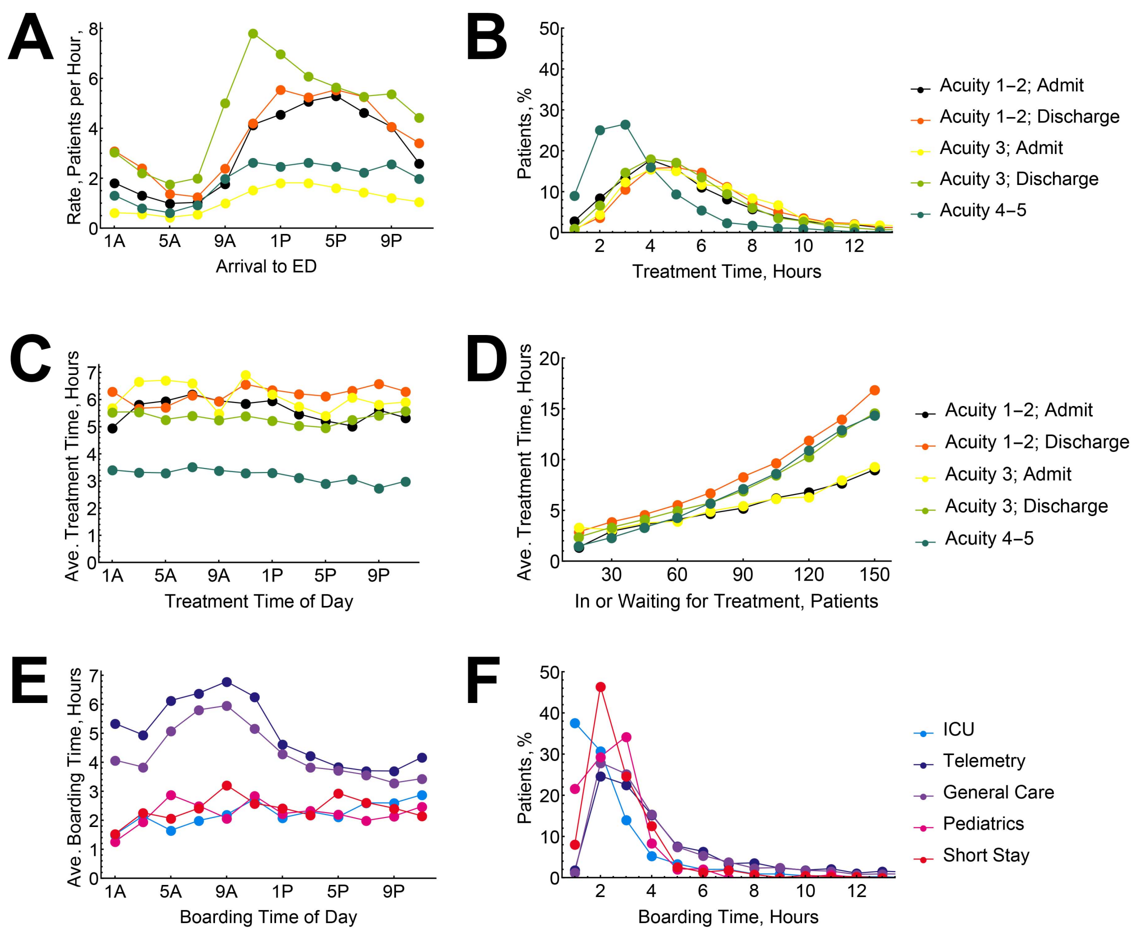

Figure 2: Processes for ED arrival, treatment, and boarding were modeled based on UMHS data. ED patients were divided into five classes based on acuity and whether they were admitted to an IU. **(A)** The ED arrival process was captured as non-homogeneous Poission process with piece-wise constant rate function that depends on time, day of the week, and class. **(B–D)** ED treatment times were modeled as lognormal random variables that depends on class and congestion, but not time. **(E–F)** ED boarding time was also modeled as a lognormal distribution that depended on where the patient was admitted and time of day.

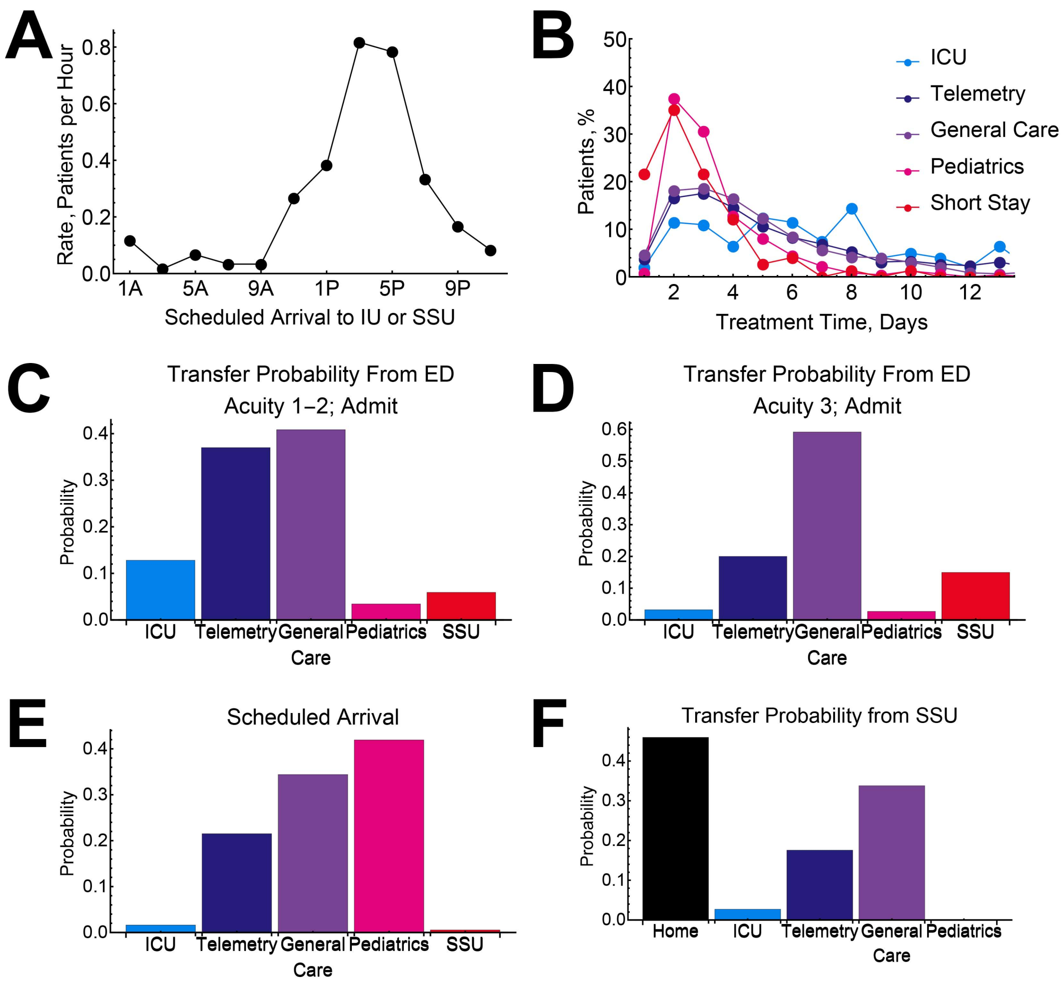

Figure 3: Process for IU arrival and treatment, along with transition probabilities for routing, were estimated from UMHS data. **(A)** The arrival process for patients that arrived initially to IU or SSU was modeled as a non-homogeneous Poisson process that depends on time. **(B)** IU/SSU treatment times were modeled as lognormal random variables that depends on the unit. **(C–D)** ED patients that were admitted would transfer to one of five inpatient units based on patient class. **(E)** Patients that arrived initially to an IU are placed in a unit based on a transition probability. **(F)** SSU patients could also transfer to another IU unit or sent home.

In accordance with current practice at UMHS, we assumed there were 119 beds for Intensive Care, 211 beds for Telemetry, 256 beds for General Medicine, and 212 for Pediatric Care. A maximum number of 69 beds were assumed for the ED, where 13 out of the 69 beds were available for new patients only between 8A and 10P, and an additional 14 out of the remaining 56 beds were available only between 11A to 10P. The number of beds for Short Stay are varied in the subsequent section. Lastly, in the ED queue, patients are prioritized based on acuity, otherwise patients are prioritized on a first-come-first-serve basis. We did not consider admission decisions for overflow in this initial model.

4.2 Application of the Model

The goal of the model is to assess the impact of SSUs on patient flow in the hospital. To illustrate the applicability of our model, we thus ran our discrete-event simulation model to examine how the number of beds in the SSU impacts the average waiting times and average bed occupancy by time of day in the ED and SSU (Figure 4).

For the discrete-event simulation, we used parameters determined by our analysis from Section 4.1 and a simulation length of 3 years, using only the the second year for analysis. Because the goal of this study was to analyze performance measures (i.e. average waiting times and bed occupancy) for a one year period, we used only the second year for analysis using the guideline that the simulation warm-up period is equal to the length of the period we wanted to simulate. Moreover, we are also evaluating average waiting times as a function of when patients enter the system. Because waiting times depend on the arrival process and congestion, we ran the simulation beyond the simulated year to ensure correct waiting times. For simplicity, we ran the simulation for another year, with the understanding that patients do not wait in queues longer than 2 days.

To assess the impact of the size of the SSU on congestion, we varied the number of SSU beds from 10 to 18, incrementing the number of beds by two each time, for a total of five scenarios. For each scenario, we calculated average hourly waiting time to ED care, average hourly bed occupancy (i.e., patients waiting in treatment) in the ED, average hourly waiting time to SSU care, and average hourly bed occupancy in the SSU by hour of day and day of week. The simulation was performed in MATLAB.

When the number of SSU beds equals 10, the system is unstable in the sense that average hourly waiting times are no longer finite. As a result, we see in Figure 4A-B that, for each day, the estimated average hourly bed occupancy in the ED and SSU equals the total number of beds in each unit: 69 in the ED and 10 in the SSU. Surprisingly, when the number of beds in the SSU increases from 10 to just 12, the system becomes stable in the sense that the average waiting times are finite.

The estimated average hourly bed occupancy in the ED generally decreases with number of SSU beds (Figure 4A). In all instances, ED occupancy follows the same sinusoidal-like, but lagged, trend as the estimated bi-hourly arrival rate (c.f. Figure 2A): relatively high occupancy in the morning that decreases to its lowest peak at 9A, and then increases to its highest peak at around 9P. The average waiting time for treatment in the ED also follow a sinusoidal-like trend: relatively high waits in the morning that decreases to its lowest peak at 6A, and the increases to its highest peak at around midnight. Lastly, average ED waiting times decreases with number of SSU beds from a peak of half an hour when the number of SSU beds is 12 at around 4A, to a trough of 0 when the number of SSU beds is 16 or 18 between 10A to 5P (Figure 4C).

In contrast, the estimated average hourly bed occupancy in the SSU stays relatively constant throughout the day (between 10 and 12 patients) and slightly increases with the number of SSU beds (Figure 4B). Moreover, the average waiting times for care in the SSU decrease with the number of SSU beds (Figure 4D). There is a very steep decrease in waiting times when the number of beds in the SSU increases from 14 beds (with a peak average wait of almost 16 hours around 10A) to 16 beds (with a peak average wait of no more than one hour). However, there is little difference in average hourly waits when the SSU has 16 or 18 beds.

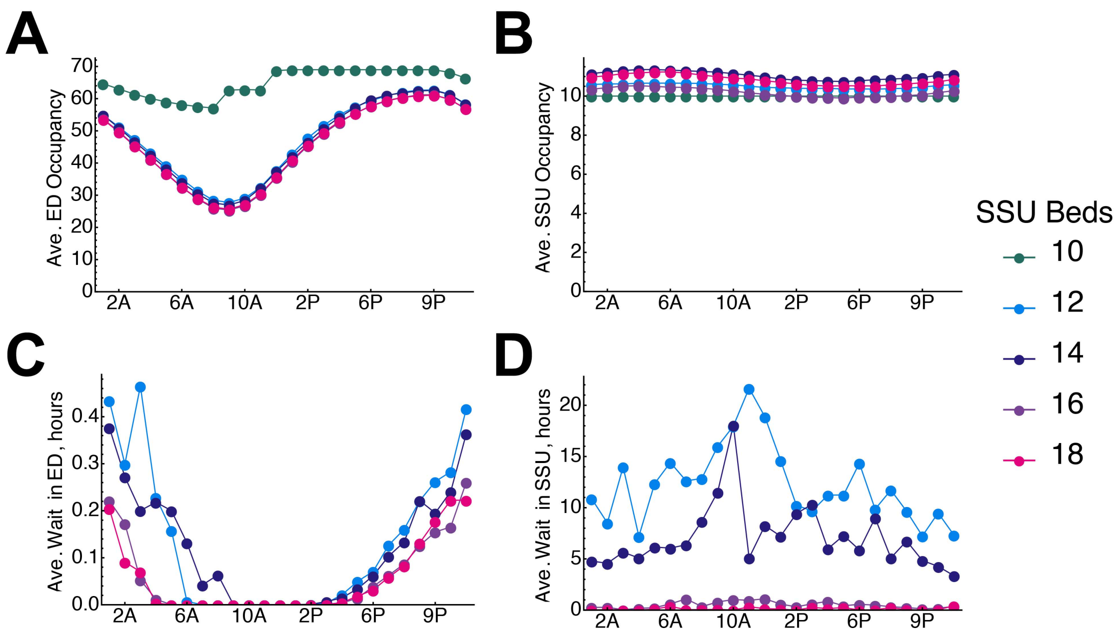

Figure 4: ED and SSU Occupancy.

5 DISCUSSION

We presented a simulation model of a hospital system as a stochastic network model, that captures many nuances of a hospital system, the most important of which is the interaction between emergency departments and inpatient units. The details of the model are based on UHMS electronic healthcare records, and is primarily motivated by our interest in broadly assessing the impact of SSUs on outcomes, efficiency, and costs of care.

Simulation models of healthcare systems have been well studied in the literature; we refer the reader to Jacobson, Hall, and Swisher (2006) and to Brailsford (2007) for systematic surveys on hospital simulation models and to Section 3.3 in Saghafian, Austin, and Traub (2015) for a summary of work focused on EDs. Our work and the work summarized in these papers underscore their importance: they allow decision-makers to analyze otherwise intractable systems and to experiment with what-if scenarios to determine feasible or appropriate system configurations without having to change the physical system.

Our paper considers a topic in healthcare simulation that has not been adequately addressed in the literature: data-driven models of hospital systems. Examples of such work include Sinreich and Marmor (2005), Marmor et al. (2009), Zeltyn et al. (2011) which focus on EDs. The authors in Sinreich and Marmor (2005), for instance, present a model of the care process in the ED based on field studies of Israeli hospitals. Their analysis allowed them to characterize ED processes in different hospitals, to propose patient types/classes (internal, surgical, orthopedic), a patient arrival processes for each patient type, and staffing levels for different medical providers. More broadly, the data allows to more rigorously identify factors that significantly contribute to patient flow in the hospital and to more precisely capture patient flow in a more realistic fashion. Moreover, it allows for identifying factors that are not significant in order to present the simplest description possible.

While all of the previous work made significant contributions in developing data-driven models, their focus is on a single unit, such as the ED. By contrast, we take a macroscopic view of the hospital system, spanning several hospital units. There is recent, relevant work in this area in Armony et al. (2015) and Shi, Chou, Dai, Ding, and Sim (2015). They study patient flow in various departments at an Israeli and Singaporean hospital. Armony et al. (2015), for example, empirically study the transfer process from ED to internal wards, and Shi et al. (2015) uses the analysis in Shi et al. (2014) to develop a simulation model to study early discharge policies that can help reduce ED boarding, widely considered a main factor contributing to ED overcrowding.

Our model also includes the ED to IU hand-off process and considers boarding, but extends previous work by including other common features observed in practice and the data (e.g. overflowing patients, priorities, blocking). We contend that these are important features to include when assessing the impact of operational decisions in a complex interacting system such as the hospital; operational changes in one unit, such as changes in the number of beds in the SSU, may have reverberations in both upstream units (e.g. ED) or downstream units (e.g. IUs). Moreover, our proposed model is a time-varying and state-dependent stochastic service network, for which no known analytic or simulation results are known.

To illustrate the former point, we generated scenarios to assess the impact of changes in the number of SSU beds on average hourly bed occupancy and average hourly waiting times for treatment in the ED and SSU over time. There are two main surprising results, with managerial implications. The first is that when the number of SSU beds equals 10, the system is unstable in the sense that average hourly waiting times are no longer finite, but can stabilized by simply increasing the number of beds from 10 to 12. This means that for the given model, hospital managers should have at least 12 beds in the SSU to avoid arbitrarily long waits. Second is that there is a significant decrease in waiting times when the number of beds in the SSU increases from 14 beds to 16 beds, but very little difference in average hourly waits when the number of SSU beds is increased from 16 beds to 18 beds. The latter is important for two reasons: first is that SSU waiting time is time spent by a patient waiting for an SSU bed, or boarding, in an ED bed, which contributes to ED overcrowding. Second is that the steep decrease in average waits when there is an increase from 14 beds to 16 beds suggests that, if the current setup has 14 beds, a manager may want to invest in additional capacity of two beds to reduce overcrowding. However, the small decrease in average waits when we increasing from 16 beds to 18 beds suggests that a manager may not want to invest in additional beds after that.

Traditional approaches to sizing hospital units are based on averages (c.g. Lovejoy and Desmond (2011)) or on sizing a given unit with the number of beds that achieves some target probability of delay (c.f. Green (2006)). Our simulation model extends these approaches by including more realistic features (e.g. stochasticity) and, in contrast to models where target probabilities are used, our model is a time-varying, state-dependent network with blocking, for which no known analytic results exist.

6 LIMITATIONS

The work presented in this paper is only a small part of a wide-ranging project, and there are several avenues for future research directions that are currently being pursued. We list two that are related to simulation. First is to better understand clinical pathways in different hospital units. For example, in the UMHS ED, an Emergency Critical Care Center has recently opened with the purpose of delivering care to the most critically ill and injured patients from the moment of arrival through the first hours of their health crisis. It is the first of its kind in the country, and a better understanding of the patient clinical pathway through this unit and how it affects flow elsewhere in the ED can help better determine whether and how it should be modeled. Similarly, we excluded operating rooms and transfers between the different IUs from our model even though operating rooms are routinely used and transfers commonly occur. Once again, a better understanding of the pathway through the operating rooms and transfers between IUs will help determine what features should be included in the model. Second is validation of our methods, continuing the limited experiment presented in Section 5, and expanding to compare them against actual hospital measurements (e.g. actual bed occupancy, ED and IU length of stay).

ACKNOWLEDGMENTS

The authors would like to acknowledge the help of Allison Cator, MD, Katrina Dawson, Jason Ham, MD, Hyeon Joo, MS, Michelle Macy, MD, and the University of Michigan Center for Healthcare Engineering and Patient Safety. GZC would like to acknowledge support from the University of Michigan's President's Postdoctoral Fellowship.

REFERENCES

Armony, M., S. Israelit, A. Mandelbaum, Y. N. Marmor, Y. Tseytlin, and G. B. Yom-Tov. 2015. "On patient flow in hospitals: A data-based queueing-science perspective". *Stochastic Systems* 5 (1): 146–194.

Brailsford, S. C. 2007. "Advances and challenges in healthcare simulation modeling: Tutorial". In *Proceedings of the 2007 Winter Simulation Conference: 40 Years! The Best is Yet to Come*, edited by R. R. Barton, B. Biller, S. G. Henderson, M.-H. Hsieh, J. Shortle, and J. D. Tew, 1436–1448. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Damiani, G., L. Pinnarelli, L. Sommella, V. Vena, P. Magrini, and W. Ricciardi. 2011. "The Short Stay Unit as a new option for hospitals: A review of the scientific literature". *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research* 17 (6): SR15.

Centers for Medicare and Medicaid Services 2011. "Medicare Benefit Policy Manual, Chapter 6: Hospital services covered under Part B". Available via https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/Downloads/bp102c06.pdf [accessed 29 November, 2016].

Green, L. 2006. "Queueing analysis in healthcare". In *Patient flow: Reducing Delay in Healthcare Delivery*, 281–307. Springer.

Jacobson, S. H., S. N. Hall, and J. R. Swisher. 2006. "Discrete-event simulation of health care systems". In *Patient flow: Reducing Delay in Healthcare Delivery*, 211–252. Springer.

Kim, S.-H., and W. Whitt. 2014a. "Are call center and hospital arrivals well modeled by nonhomogeneous Poisson processes?". *Manufacturing & Service Operations Management* 16 (3): 464–480.

Kim, S.-H., and W. Whitt. 2014b. "Choosing arrival process models for service systems: Tests of a nonhomogeneous Poisson process". *Naval Research Logistics (NRL)* 61 (1): 66–90.

Lewis, P. A. 1965. "Some results on tests for Poisson processes". *Biometrika* 52 (1-2): 67–77.

Lovejoy, W. S., and J. S. Desmond. 2011. "Little's Law flow analysis of observation unit impact and sizing". *Academic Emergency Medicine* 18 (2): 183–189.

Marmor, Y. N., S. Wasserkrug, S. Zeltyn, Y. Mesika, O. Greenshpan, B. Carmeli, A. Shtub, and A. Mandelbaum. 2009. "Toward simulation-based real-time decision-support systems for emergency departments". In *Proceedings of the 2009 Winter Simulation Conference*, edited by A. Dunkin, R. R. Hill, R. G. Ingalls, B. Johansson, and M. D. Rossetti, 2042–2053. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Noel-Miller, C., and K. Lind. 2015. "Is observation status substituting for hospital readmission". *Health Affairs Blog*.

Sabbatini, A. K., B. K. Nallamothu, and K. E. Kocher. 2014. "Reducing variation in hospital admissions from the emergency department for low-mortality conditions may produce savings". *Health Affairs* 33 (9): 1655–1663.

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

Saghafian, S., W. J. Hopp, M. P. Van Oyen, J. S. Desmond, and S. L. Kronick. 2012. "Patient streaming as a mechanism for improving responsiveness in emergency departments". *Operations Research* 60 (5): 1080–1097.

Saghafian, S., W. J. Hopp, M. P. Van Oyen, J. S. Desmond, and S. L. Kronick. 2014. "Complexity-based triage: A tool for improving patient safety and operational efficiency". *Manufacturing and Service Operations Management*. to appear.

Shi, P., M. C. Chou, J. Dai, D. Ding, and J. Sim. 2015. "Models and insights for hospital inpatient operations: Time-dependent ed boarding time". *Management Science* 62 (1): 1–28.

Shi, P., J. Dai, D. Ding, S. ANG, M. Chou, X. Jin, and J. Sim. 2014. "Patient flow from emergency department to inpatient wards: Empirical observations from a Singaporean hospital". *Mabel and Jin, Xin and Sim, Joe, Patient Flow from Emergency Department to Inpatient Wards: Empirical Observations from a Singaporean Hospital (October 30, 2014)*.

Sinreich, D., and Y. Marmor. 2005. "Emergency department operations: the basis for developing a simulation tool". *IIE Transactions* 37 (3): 233–245.

Zeltyn, S., Y. N. Marmor, A. Mandelbaum, B. Carmeli, O. Greenshpan, Y. Mesika, S. Wasserkrug, P. Vortman, A. Shtub, T. Lauterman et al. 2011. "Simulation-based models of emergency departments:: Operational, tactical, and strategic staffing". *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 21 (4): 24.

Zuckerman, R. B., S. H. Sheingold, E. J. Orav, J. Ruhter, and A. M. Epstein. 2016. "Readmissions, observation, and the hospital readmissions reduction program". *New England Journal of Medicine* 374 (16): 1543–1551.

AUTHOR BIOGRAPHIES

JOHN T. CHAVIS III is a PhD student in the Center for Applied Mathematics at Cornell University. His research interests are in Applied Probability with a focus on Markov decision processes, queueing theory, and simulation. His email address is jc278@cornell.edu.

AMY L. COCHRAN is a T.H Hildebrandt Research Assistant Professor (postdoc) in Mathematics at the University of Michigan. Her research interests are in mathematical biology. In particular, she uses stochastic processes, statistics, and differential equations to tackle problems in psychiatry. Her email address is cochraam@umich.edu.

KEITH E. KOCHER, MD MPH, is an Assistant Professor in the Department of Emergency Medicine at the University of Michigan Medical School. He is an emergency physician and health services researcher who studies the delivery and performance of emergency and acute care. His email address is kkocher@umich.edu.

VALERIE N. WASHINGTON is an Industrial and Systems Engineering undergraduate student at Kennesaw State University. She is a research assistant in projects encompassing healthcare and fitness tracking technologies. Her email address is vwashin8@students.kennesaw.edu.

GABRIEL ZAYAS-CABÁN is a President's Postdoctoral Research Fellow in the Center for Healthcare Engineering and Patient Safety in the Industrial and Operations Engineering Department at the University of Michigan. His present research specialty is on stochastic modeling and optimization with a focus on healthcare operations, medical decision-making, and healthcare policy. His email address is gzayasca@umich.edu.