Simulated co-location of patients admitted to an inpatient internal medicine teaching unit: Potential impacts on efficiency and physician-nurse collaboration

Blair L. Bilodeau[1], David A. Stanford[1], Mark Goldszmidt[2], Andrew Appleton[2]

[1]Department of Statistical & Actuarial Sciences, Western University, London, Canada.
Email: bbilodea@uwo.ca, stanford@stats.uwo.ca;
[2]Department of Medicine, Schulich School of Medicine & Dentistry, Western University, London, Canada. Email: Mark.Goldszmidt@schulich.uwo.ca, Andrew.Appleton@lhsc.on.ca

ARTICLE HISTORY
Compiled May 25, 2019

This article has been accepted for publication at INFOR: Information Systems and Operational Research. Once live, please view at https://doi.org/10.1080/03155986.2019.1624475.

ABSTRACT
Effective communication between nurses and physician teams on the internal medicine unit is crucial for high quality, safe, and efficient patient care. In our hospital of interest, a large academic health sciences centre, the physical layout of the unit, current admission process, and the presence of three separate physician teams contribute to uneven workload and communication barriers. We address this by physically co-locating each physician teams' patients so as to facilitate physician-nurse collaboration, and more evenly distribute workload across all three teams. Based upon one year of real-world data, we developed a simulation model of inpatient flow through the internal medicine unit and determined the impact of two proposed changes: co-locating each team's patients, and new admission rules for how patients are assigned to those teams. Under the new arrangement, each physician team would interact with roughly half the number of nurses, and nurses in turn would have fewer individual team members with whom to communicate, thereby improving effective communication and increasing time for direct patient care for both physicians and nurses.

KEYWORDS
patient flow; patient co-location; simulation; physician-nurse collaboration

1. Introduction

A common problem for physician teams working in inpatient settings is the time it takes to collaborate with the nursing staff involved in the care of their patients. When a particular team's patient load is spread over a large number of nurses, this is more time consuming, and thus negatively impacts the amount of time available for direct patient care. This problem arises naturally as a function of the current bed assignment policy at University Hospital, an academic health sciences centre in London, Ontario. Concerns over these issues motivated the work for this paper, which considers changes to the physician team and bed assignment policies in the hope that improving the efficiency of the system would translate into improved patient care.

We sought to develop a patient flow procedure that would retain current staffing levels while focussing on physician and nurse efficiencies simultaneously. A discrete event simulation model of the proposed changes was developed using the open source Python programming language (see Python 2017). By physically co-locating patients and employing a novel method for assigning patients to physician teams, we report a minimally disruptive administrative strategy that is expected to result in better patient care.

2. Literature review

The sheer volume of papers on simulation in health care is so large that one could, perhaps, suggest that it would be appropriate to have a survey paper about the survey papers in the field. Surveys include England and Roberts (1978), Lehaney and Hlupic (1995), Jun et al. (1999) and Fone et al. (2003). These last two are considered by Günal and Pidd (2010) to be "more comprehensive and systematic than their predecessors" (p. 43). Günal and Pidd's own review considers many dozens of papers up to 2008; and (among other things) finds that two common tendencies with the simulation models they review are that they are frequently unit specific, focussing on only one aspect of the hospital, and facility specific, making it hard for the conclusions to be extended to sites other than the one where the study was carried out. They also comment on the lack of implementation of simulation study conclusions as a continuing problem, more than 25 years after the problem was noted in Wilson (1981).

Jacobson et al. (2006), as a book chapter, has the space to provide greater depth in its review of the literature during the period 1965 to 2004, with the dominant bulk of the review focussing on papers from the latter 20 years of that period. Jacobson et al. (2006) complement the unit-specific view with respect to patient flow and related issues, with a detailed section on health care asset allocation, including issues in sizing and planning, facility planning, and nurse / physician scheduling.

Brailsford et al. (2009) classify 342 papers pertaining to analytical approaches in health care from the period 1952-2007, with the bulk coming after 1990. Almost two-thirds are classified in the fields of statistical analyses and modelling, with others belonging primarily to the fields of mathematical programming, qualitative studies, simulations, and a pool of other areas. About one sixth were simulation studies.

The references in Hu et al. (2018) provide many examples of how simulation techniques have aided the study of patient flow in hospital emergency departments. Most commonly, the goal is to reduce patient length of stay and waiting time by adjusting how patients move through the ward (Klein and Reinhardt 2012; Powell et al. 2012; Rasheed et al. 2012). Shenoy et al. (2018) consider how to assign patients efficiently to beds given infection restrictions, while de Bruin et al. (2007) look at optimizing bed allocation to explore bottlenecks in patient flow. Best et al. (2014) and Rossetti et al. (1999) investigate optimizing physician staffing based on patient occupancy levels, while Agor et al. (2016) do so through the lens of a novel proxy metric for physician workload. Sarno and Nenni (2016), Siddiqui et al. (2017), and Sundaramoorthi et al. (2010) all look at variations of adjusting nurse staffing levels to standardize the workload distribution. Acar and Butt (2016) also look at nursing assignments, but with specific consideration for the distance nurses travelled between patients.

Much closer to the present work is Mandelbaum et al. (2012), which addresses the issue of fair routing of patients. The authors employ a specialized queueing model with heterogeneous servers to address the reality that differing individuals or teams render treatment at different rates. Their key contribution is the introduction of the Randomized Most Idle (RMI) policy for assigning customers (patients) to available server pools (wards) on the basis of the number of idle servers in each group. This has the advantage over a Longest Idle Server First (LISF) policy in that the latter requires detailed information that is not usually available to decision makers in real time.

3. Current process

The internal medicine inpatient teaching unit – hereafter referred to as the unit – treats large volumes of patients with acute medical illness requiring hospital admission. Most patients are admitted to internal medicine through the emergency department, although some are transferred from other inpatient services. Most also suffer from multimorbidity and thus are considered to be among the most complex patients to care for in the hospital. This section details the important characteristics of the unit along with how patients flow from admission to discharge.

3.1. Unit layout

The unit has 72 total beds; 60 on the fourth floor and 12 on the sixth floor. The fourth floor is the main internal medicine unit and is the preferred locale for treating high acuity patients. It has 10 private rooms (single bed per room), 18 semi-private rooms (2 beds per room), and 4 ward rooms (3 or 4 beds per room). The sixth floor has 5 private rooms and 4 semi-private rooms (although only 7 semi-private beds, as shown in Figure 1). In addition, there are 6 'decant' beds on the sixth floor. Decant beds are intended to offload patients temporarily from the emergency department during the day while they wait for one of the 72 beds to open. In reality decant beds often remain open overnight, with one or two patients present.

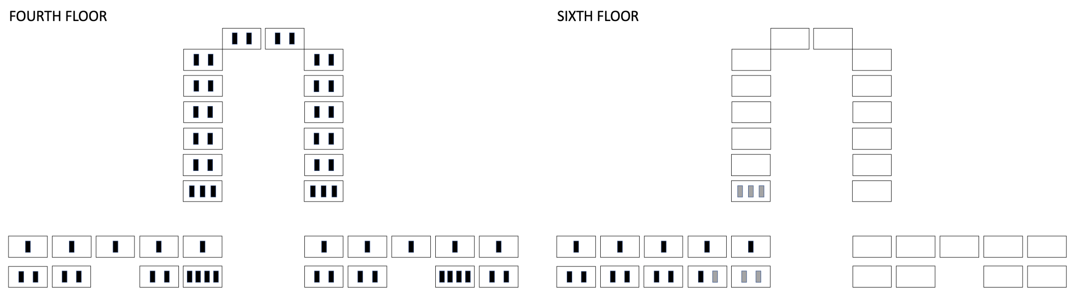

Figure 1. The internal medicine inpatient teaching unit beds on the fourth and sixth floors of University Hospital. Grey beds indicate decant.

3.2. Staffing levels

The unit is served by three distinct physician teams. This is done partly to ensure equal workload and breadth of experience for medical students and resident doctors, who complete the initial medical assessment of new patients. On weekdays, from 8 a.m. to 5 p.m., each team consists of one attending physician, one third year resident, one second year resident, three to four first year residents, and three medical students. During the same hours on weekends, each team has one attending physician, one second or third year resident, and one medical student or first year resident. On all days of the week, from 5 p.m. to 8 a.m., the on-call team includes one first year resident or medical student from each of the three teams, plus one second or third year senior resident who oversees the other three more junior trainees.

Nursing staff provide most of the direct patient care. Registered nurses (RNs) may be assigned to any patient on the unit, while registered practical nurses (RPNs) typically care for lower acuity patients; these RPNs cannot perform some procedures. Each 24-hour period consists of a day shift from 7 a.m. to 7 p.m. and a night shift from 7 p.m. to 7 a.m. Prior to each shift the charge nurse assigns the incoming nurses to their patients. This is done using a subjective process with consideration of skill level, years of experience, and personality. Table 1 outlines the number of RNs and RPNs on the unit at all times. Regardless of time of day, a nurse will always be assigned patients in nearly adjacent rooms. As the unit has not reported any difficulty in staffing for these shifts, we have not addressed staffing matters in this study.

Table 1. Nursing staff levels by floor and time of day.

		7a.m. - 7p.m.	**7p.m. - 7a.m.**
Fourth Floor	RNs	13	9
	RPNs	2	1
Sixth Floor	RNs	3	2
	RPNs	0	0

As seen in Figure 1, the fourth floor of the unit comprises three hallways. Each of the three physician teams may have patients in any hallway and on both the fourth and sixth floors. How these beds are used among the teams is one of the issues we address in the analysis which follows.

3.3. Unit statistics

University Hospital provided a dataset containing summary statistics (mainly means and some variances) of various performance metrics for all patients who were admitted to internal medicine between March 2016 and April 2017. These are routinely collected administrative data with no patient-identifying information.

The admission data which the hospital provided was limited to the annual number of patients admitted to internal medicine, along with the percentage of patients admitted during the six different time slots. Table 2 contains the corresponding rates calculated from these percentages. As no information on seasonality or day to day variability was provided, we consider the average admission rates applicable for each day of the week.

A patient's origin is defined as where they are in the hospital when the decision to admit them to internal medicine is made. Most commonly, this is the emergency department, although patients can be admitted from elsewhere in the hospital. Additionally, patients from other units who are occupying a unit bed off-service are not actually admitted to internal medicine. While these off-service patients impact internal medicine patients and thus affect performance metrics indirectly, their values are not directly included in the calculation of these metrics. The status of a patient is defined to be the minimally required type of bed they can use. That is, a ward patient

Table 2. Rate of patients admitted to internal medicine per hour.

Time Period	Admission Rate
00:00 – 03:59	0.48
04:00 – 07:59	0.24
08:00 – 11:59	0.63
12:00 – 15:59	0.83
16:00 – 19:59	0.83
20:00 – 23:59	0.67

can use a private or semi-private bed, but a private patient must use a private bed. This is approximated by the number of patients admitted with *Clostridium difficile*, Methicillin-Resistant *Staphylococcus aureus* infection, or droplet precautions. There was no data available on patients' insurance status to further determine room preferences. Finally, the acuity of a patient is classified as either high, moderate, or low, following the guidelines of the College of Nurses of Ontario (2018). This acuity data was more difficult to obtain and consequently is only from a sample of 7 days worth of patients and expert opinion.

The data does not exactly specify these proportions, so we obtained additional input from the unit physicians. For instance, the number of patients admitted from the hospital's ICU was overrepresented in the data as it also counted medicine patients who were briefly transferred there. Thus, the percentages in Tables 3–5 are a hybrid of the given data and expert opinion.

Table 3. Percentage of patients requiring each bed type. Referred to as *status*.

Bed Type	Patient Percentage
Private	20%
Semi-Private	25%
Ward	55%

Table 4. Percentage of patients admitted from each location. Referred to as *origin*.

Location of Admission	Patient Percentage
Emergency Department	81.0%
Non-Medicine	13.7%
Intensive Care Unit	4.6%
Other	0.7%

A patient's length of stay (LOS) is defined to be the time from when they are admitted to internal medicine until they are discharged, regardless of when they actually enter a medicine bed. The data recorded for LOS was in whole day units, with a mean of 5.3 and standard deviation of 5.7. Additional data specified the percentage of discharge orders sent for each hour of the day, which is displayed in Table 6. The mean time between a discharge order being placed and the patient being discharged was observed to be 2.1 hours.

3.4. Unit procedures and operations

Regardless of time of day, all referrals from the emergency room are received, triaged, and seen by a designated senior medical resident. Between 8 a.m. and approximately

Table 5. Percentage of patients admitted with each acuity level. Referred to as *acuity*.

Acuity Level	Patient Percentage
High	5%
Moderate	60%
Low	35%

Table 6. Percentage of patients discharged during each hour of the day.

Hour	00	01	02	03	04	05	06	07
Patient Percentage	0.00	0.00	0.00	0.00	0.00	0.09	0.09	2.97
Hour	08	09	10	11	12	13	14	15
Patient Percentage	4.83	7.81	17.47	23.42	10.41	12.27	8.92	4.46
Hour	16	17	18	19	20	21	22	23
Patient Percentage	3.72	1.12	1.49	0.37	0.28	0.09	0.09	0.09

6 p.m., they will see the patients on their own and arrange for their admission to one of the teams. After 6 p.m., new referrals will be assigned to one of the three on-call junior trainees and subsequently admitted to the physician team they are assigned to. For the most part, this current process involves evenly cycling so that every third patient is assigned to the same junior trainee and therefore the same physician team, which we refer to as the "rolling basis" method. In practice there are deviations to this protocol, due to information loss at shift changes and subjective decision-making on the part of the designated senior medical resident.

Once the senior medical resident has assigned the patient to a physician team and requested admission to internal medicine, the bed assignment process is handled administratively. Certain criteria at the time of admission will impact the type of room a patient is assigned. For instance, patients with a *Clostridium difficile* infection require a private room to avoid spreading the infection. However, due to volume demands, private rooms are often occupied by patients who do not require them. If there is no suitable available bed, the patient may be placed in a decant bed if there is one available during daytime hours. Otherwise, they may be placed in an off-service bed elsewhere in the hospital. If none of these options are available, the patient will remain in the emergency department until an appropriate unit bed becomes available.

Unit beds are vacated when a patient is discharged, transferred, or has died. Once vacated, the patient area and bed are cleaned. The cleaning process can take up to two hours depending on whether the patient had a contagious disease that can be acquired environmentally. The bed is then available for the next suitable patient who is waiting, or will remain empty until a suitable patient is admitted. Patients waiting for a bed are prioritized based on where they are currently located. Patients in the emergency department or intensive care unit receive top priority, followed by patients in decant beds, then patients who have been placed off-service, and finally patients admitted to other services whose units are fully occupied.

4. Issues and proposed changes

As noted earlier, one source of inefficiency for internal medicine teams is the number of nurses with whom they interact while performing rounds on their patients. While each nurse will be assigned four patients during the day, this assignment is done at present without consideration for which physician team is caring for the patients involved. Thus, a nurse could have patients from each of the three teams, thereby imposing an additional burden to track the physicians or medical students in charge of their patients. Instead of having one conversation about three or four patients, each nurse might have a separate conversation with each team from which they are assigned patients. Limiting the number of nurses each physician team interacts with by assigning them to a greater proportion of a single team's patients would reduce this consultation overhead. Dobson et al. (2009) address the impacts of communication barriers, albeit in a primary care setting.

The number of nurses a team has caring for their patients is influenced by how many patients the team has, which can vary significantly. To account for potential noise introduced by census variations, we define the metric patients per nurse (PPN) for each team. This is calculated four times a day (immediately prior to and immediately following both nursing shift changes) by taking the number of patients a team has in the 72 medicine beds, and dividing by the number of nurses assigned to that team's patients. This metric is used as a proxy for nursing efficiency, where a high PPN value is preferable, signalling a lower number of nurses for the team to collaborate with. At shift change, the patients sharing a nurse will change, which will affect the PPN value. This change is likely not going to be large, but calculating before and after allows us to evaluate how PPN evolves over the shift without the added noise from a shift change.

Another objective for the physician teams is to strive to ensure that residents and medical students see a comparable number of patients each day for educational purposes. The practice of assigning teams on a rolling basis ensures that each team sees approximately the same number of new patients in a week. However, the large variability in patient length of stay means that the distribution of patients among the three teams at a given time can be very uneven. This means that the total amount of daily work, driven by the total number of inpatients, is not evenly distributed across teams despite their equal staffing levels. In an effort to address this, we define the metric team census variance (TCV). In the dataset provided, the number of patients on each team is only recorded in the morning of each day, rather than at the midnight census as with much of the other data. For consistency, the model also includes a morning census to record TCV. Each instance of this metric is simply the variance of the three team census values in the morning, giving a single number for each day; these instances can then be averaged over every day to see the average TCV for a given year.

From both a patient and hospital flow perspective, it is desirable to assign the patient to the first available bed. From the team's perspective, it is desirable to assign the patient to the team with the lowest census, to ensure that each team gets a comparable workload each day. Fortunately, these goals are positively correlated. The rule we have implemented, which we refer to as the co-location method, is as follows. Each of the 72 medicine beds is assigned to a team as per Figure 2. Primarily, we assign a new patient to the first available bed. If multiple beds are available, the patient goes to the one associated with the team that has the lowest census at that time. When no bed is available upon admission, the patient will be initially assigned a team using the lowest census method. Once a bed becomes available, the patient is moved and reassigned, if necessary, to the team associated with this bed. Importantly, this rule does not change the charge nurse procedure for scheduling nurses or the physician team staff levels.

Our research goal is to determine if implementing the co-location method will yield an increase in PPN values and a decrease in TCV while not negatively impacting the congestion of the unit or the waiting time in the emergency department. By still allowing patients to take the first available bed we expect to see no increase in waiting time or bed utilization rates. Since each team's patients will now be in adjacent rooms, they are much more likely to be sharing a nurse (as nurses are assigned a patient based on this criteria). This will directly increase the team's PPN value, which we hypothesize will improve collaboration efficiency for physicians and nurses.

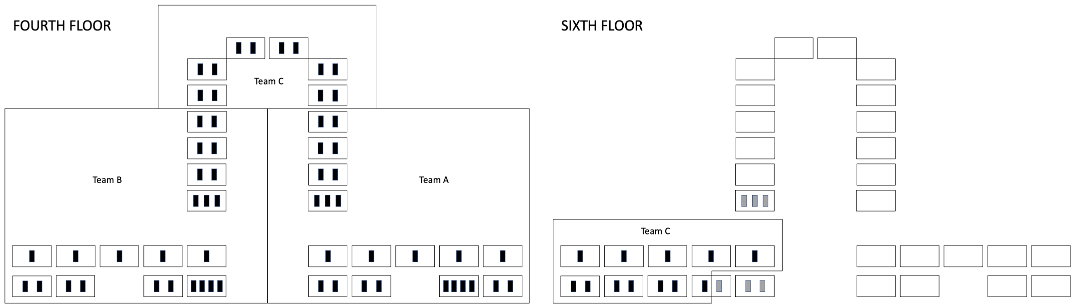

Figure 2. Proposed team assignment by bed for the fourth and sixth floors at University Hospital.

5. Simulation model

We chose to implement this simulation model in Python both for speed of computation and due to the first author's familiarity with the language. We investigated Python's simulation library SimPy, as well as the commercial software Simul8, but ultimately decided that these were unnecessary for what we required, and so the simulation logic was written from scratch. The only libraries used were NumPy for random number generation and Pandas for data storage and manipulation.

At the core of discrete event simulation is the calendar of events, which contains future actions to be simulated and is continuously updated as events are processed and new ones are generated. For this model, key events considered in the calendar are patient admissions, patient discharges, a bed being cleaned, and the start of a nursing shift. These are outlined in Figure 3, and follow a similar basic structure to the one described by Hamrock et al. (2013).

When running our simulation, the model begins from an empty state; none of the medicine beds are occupied and nobody is waiting in the emergency department. It is then run for 26 weeks of warm-up, followed by 52 weeks of data collection. This process is repeated 20 times, so that 20 independent years of data have been collected.

This warm-up period is consistent with the length prescribed by Welch's method (Law 2007), which is determined by plotting a smoothed moving average of multiple runs and visually evaluating a typical amount of time until convergence. Law also suggests an iterative algorithm to obtain a number of runs that results in error below a certain threshold for key metrics of interest. We deemed 20 runs to be an acceptable balance between this error level and time constraints for running the simulation.

Since the dataset provided which we described in Section 3.3 only contains aggregate information, determining which distributions to use for generating the random values was a process of intuition and observing how well the model output matched the expected results. An admission event occurs when a patient inter-admission time is generated using a non-homogeneous Poisson process, which is a common modelling

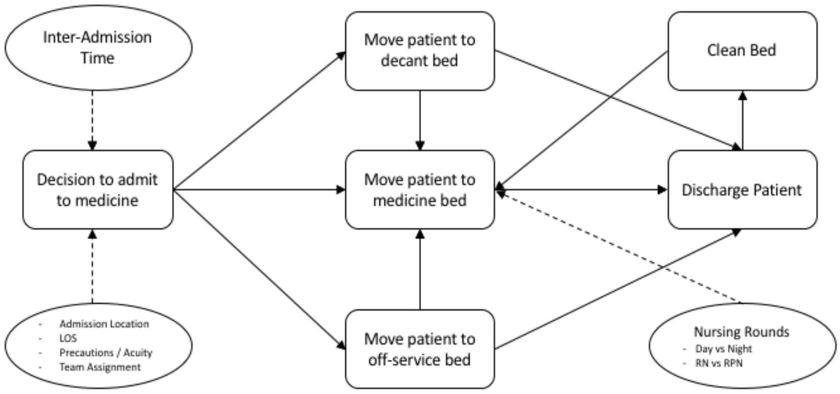

Figure 3. Flow chart for modelling patient flow through the internal medicine unit at University Hospital.

assumption for rare processes occurring in a relatively large population. Status, origin, and acuity provide further context for a specific patient's care requirements and determine how they will be moved through the model. The multinomial distribution was the logical choice for classifying patients for each of these three parameters. In reality, a patient's status and acuity may change multiple times during their treatment. However, due to the lack of data on these changes, they have been treated as static for the entire length of stay after being initialized at time of admission.

Upon being generated in the simulation, a patient is assigned to one of the three physician teams unless they are a non-medicine patient being placed off-service into an internal medicine bed. Team assignment, along with medicine bed assignment, is done according to the version of the simulation being run. For admission occurring between 7 a.m. and 7 p.m., when there is no bed available, the admitted patient will be placed in one of the sixth floor decant beds if possible. If this is also unavailable, a patient who has been admitted from the emergency department or intensive care unit may be placed off-service in an inpatient bed somewhere else in the hospital. After all of these options have been exhausted, the patient will remain in the emergency department until either a decant bed or internal medicine bed becomes available. Note that all patient transportation times are considered negligible in the model.

Once a patient has been admitted to internal medicine, their treatment begins immediately regardless of their location (including the emergency department). For simplicity, we assume this doesn't affect overall treatment time, so the patient's total length of stay in the model is determined at the time of the patient admission event. This is calculated as the sum of the geometrically-distributed whole number of days plus the fractional part of a day based on the hour when the discharge order is placed. Table 6 is used as the basis for the determination of the fractional part's duration. At the instant that their length of stay is completed, a patient discharge event is triggered. For the purposes of this model, all discharge events are treated in the same fashion regardless of whether they are discharges, transfers, or deaths.

When assigning nurses to patients prior to each shift, the charge nurse will consider factors such as the compatibility of patient and nurse personalities. However, such aspects are too subjective for inclusion in the model, so our simulation only considers whether a nurse is an RN or an RPN when determining the nurse assignment for each shift. Consequently, the model only prescribes which beds are sharing a nurse and the type of nurse, without indicating which specific nurse is caring for these beds. Since the additional specification does not affect PPN or TCV, the model adequately reflects the essential aspects of the nurse assignment process.

6. Model verification

There were many variables provided in the dataset used to quantify the model output and verify that the simulation adequately approximates reality for our purposes. This section highlights a few of these key metrics that are indicative of the congestion experienced in the unit. The utilization metrics are defined to be the percentage of the 72 internal medicine beds that are occupied. For floor utilization, this considers all patients in beds, while for medicine utilization only beds with an internal medicine patient in them (and thus not an off-service patient) are counted.

Table 7. Performance metrics for the reference simulation compared to the values observed in the dataset provided.

Performance Metric	Observed Expected Value	Simulation 95% CI
Avg. Waiting Time* (Hrs.)	6.4	(6.7, 7.4)
Avg. No. of Admitted Patients Waiting	3.4	(3.0, 3.3)
Avg. Floor Utilization Rate	94.8%	(94.3%, 95.2%)
Avg. Medicine Utilization Rate	83.5%	(82.1%, 83.2%)

*Defined to be the time it takes for a patient to leave the emergency department after admission to internal medicine.

In the performance metrics of Table 7, the average time to leave the emergency department has been calculated as the usual average over all the patients who were admitted to medicine during the study period. However, the occupancy and utilization rates have been based upon information that is only available from the daily midnight census. Therefore, daily snapshots at midnight were taken from the simulation runs so as to provide a meaningful comparison.

Clearly, the simulation doesn't capture everything that is relevant to the operation of the unit, as three out of four confidence intervals do not contain the expected value we have observed. This is not unexpected due to the limited information on parameter distributions provided in the dataset, so we are more interested in verifying the model against these parameters. We notice that the parameter estimates are each individually in the right "ballpark" as those from the dataset, and if the intervals were loosened to the 99.5% level then all the observed values would be captured. Thus, we have concluded that the model is useful for decision making purposes.

Since the model is not able to recreate the values from the dataset, we cannot infer that any changes to the model would output the same nominal values that would be observed from such a change in real life. This is compounded by the fact that the given data is both coarse and only reports on a year of operation, which means it is likely that the values which we come up with in the simulation will be biased relative to the long run averages seen on the unit. Instead, we compare the proposed changes with the reference simulation to see if the two simulations are statistically different from each other. From this, we conclude the direction of the impact on performance metrics from introducing co-location in reality, mindful that the precise values we obtain are not likely to be replicated in real life.

7. Results

Below, we present the results of testing the co-location method relative to the reference simulation rather than the original dataset, which contained limited amounts of information. Additionally, we use a common random numbers approach, which means for each run we test the same sequence of patient admissions under the two bed/team assignment policies. The performance of the co-location method was assessed using the following metrics: emergency department waiting time, bed occupancy rate, PPN, and TCV. Ultimately, we obtained results that confirmed improvements in PPN and census variance without adversely impacting emergency department waiting time or bed occupancy rate.

7.1. Congestion

Table 8 compares the performance of the co-location method to the reference simulation for waiting time and occupancy level. We wish to determine if there is a statistical difference between these two patient assignment methods for each metric. Since the same sequence of inter-admission times is used to compare the methods for an individual run, and averages are approximately normally distributed, it is appropriate to use a paired t-test to determine whether the average difference is non-zero. As the p values indicate, we see little evidence to reject the null hypothesis, which is that there is no difference.

Table 8. Performance metrics for the co-location method compared to the values from the reference simulation. P values refer to a one-tailed paired t-test of $H_A : \mu_{\text{co-location}} > \mu_{\text{reference}}$.

Performance Metric	Reference Simulation	Co-location 95% CI	P Value
Avg. Waiting Time (Hrs.)	7.1	$(6.8, 7.4)$	> 0.05
Avg. No. of Admitted Patients Waiting	3.1	$(3.0, 3.3)$	> 0.05
Avg. Floor Utilization Rate	94.7%	$(94.3\%, 95.1\%)$	> 0.05
Avg. Medicine Utilization Rate	82.7%	$(82.2\%, 83.1\%)$	> 0.05

7.2. Patients per nurse

Co-location leads to a significant increase in PPN values for each team, as seen in Table 9. For each team and at all times, a paired t-test provides evidence to reject the null hypothesis of no difference between average PPN values with a significance level < 0.01. We note that Team C now has a lower PPN than Teams A and B at night, which is caused by Team C beds being split between floors four and six.

Additionally, while one might hope that the optimal PPN for each team is 4.0 during the day and 6.0 at night, we determine the actual optimal values to be these nominal nursing staff levels multiplied by the average utilization rate. This adjustment accounts for off-service patients taking up medicine beds.

7.3. Team census variance

A version of the simulation identical to the reference, except that it implemented lowest census team assignment, was used to quantify the effect on TCV; four graphs pertaining to a single indicative run are plotted in Figure 4. These demonstrate that using the lowest census will greatly reduce the TCV, but consequently increase the

Table 9. PPN values for the co-location method compared to the values from the reference simulation.

		Start Day	**End Day**	**Start Night**	**End Night**
Reference Simulation		1.5	1.5	1.8	1.9
Co-location	Team A	3.2	3.1	4.4	4.7
	Team B	3.2	3.1	4.4	4.7
	Team C	3.3	3.1	3.9	4.1
Optimal PPN		3.4	3.4	5.1	5.1

variance for the number of new patients admitted per team. This is expected, since if one team is discharging a lot of patients they are also going to see more new patients. However, we feel it is more important to balance overall workload as opposed to the number of new patients seen.

The impact of the co-location method on TCV compared to the mean reference value of 6.4 is a reduction to 0.4. A paired t-test provides evidence to reject the null hypothesis that the average team census variance is the same between the two methods at a significance level < 0.01.

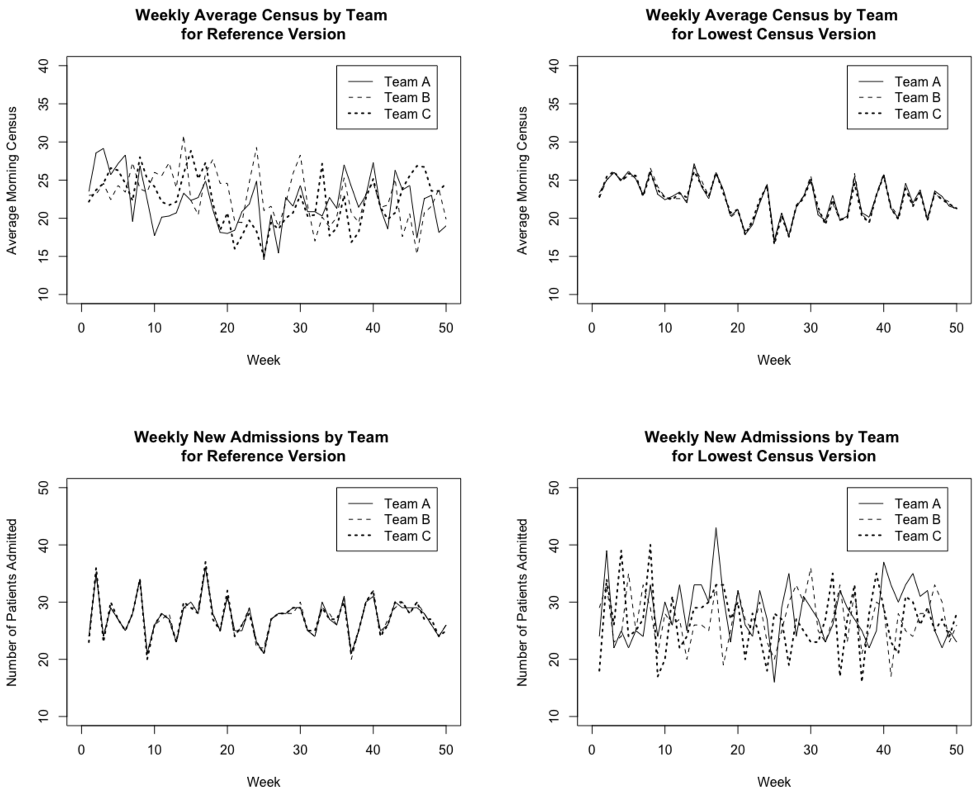

Figure 4. Comparison of rolling basis vs. lowest census team assignment methods for each team from an indicative simulation run.

One drawback of the co-location method is that patients may change teams when they are moved into a medicine bed if one was not available at admission time. Our simulation shows that team reassignment in this fashion is required one-third of the time. Ideally, such team changes would occur early in the patient's stay. This is generally the case; our simulation results show 12% of patients change between 12 and 24 hours after admission, and an additional 6% change after more than 24 hours.

8. Discussion

To realize efficiency improvements without impacting the other system performance metrics, our simulation changes the current patient assignment system in two major ways. First, we assign each bed a team by geographical groupings according to the plan presented in Figure 2, so that each team has a specified area for their patients. Second, if a medicine bed is available, we assign a new patient to the corresponding team, whereas if none is available the patient is initially assigned to the team with the lowest census. The patient is then reassigned to another team if the first available internal medicine bed is associated with another team.

The University Hospital executive leadership has commited to implementing the co-location method. At this time, the fourth author has submitted a grant proposal to fund the medical researchers needed to oversee this. For an implementation to be successful, the model should be reconciled with the following constraints.

Ideally, in the context studied, the team changes should be made prior to assigning a junior trainee to see the patient and certainly prior to the patient meeting with the attending physician. For both patients and physicians, continuity of care needs to be preserved. In other settings, for example those using night float systems where the on-call juniors are not attached to a specific team, this would be less of an issue (see Bricker and Markert 2010). In situations where the implementation must sacrifice either continuity or bed availability, we recommend developing a policy with the physician teams and testing it over time. One option that prioritizes continuity would be to make daily corrections based on admissions taking place during the daytime when a senior medical resident is seeing patients independently, as the designated senior medical resident is not attached to a team.

For the purposes of this paper, the main aspects of the unit have been captured to create a reasonably accurate simulation which can be run easily in the Python programming language. The major limitation that affects this model is that the distributions for patient inter-arrival times and patient length of stay are unknown. Nonetheless, it is our perspective that a more refined model would come to the same conclusions.

Funding

This work was supported by the first author's NSERC USRA and the second author's NSERC Discovery grant.

References

Acar I, Butt S. 2016. Modeling nurse-patient assignments considering patient acuity and travel distance metrics. J Biomed Inform. 64:192–206.

Agor J, McKenzie K, Ozaltin O, Mayorga M, Parikh RS, Huddleston J. 2016. Simulation of triaging patients into an internal medicine department to validate the use of an optimization based workload score. In: Roeder TMK, Frazier PI, Szechtman R, Zhou E, Huschka T, Chick SE, editors. WSC 2016. Proceedings of the 2016 Winter Simulation Conference; 2016 Dec 11–14; Washington, DC. Washington (DC); IEEE. p. 3708-3709.

Best AM, Dixon AC, Kelton WD, Lindsell CJ, Ward MJ. 2014. Using discrete event computer simulation to improve patient flow in a Ghanaian acute care hospital. Am J Emerg Med. 32(8):917–922.

Brailsford SC, Harper PR, Patel B, Pitt M. 2009. An analysis of the academic literature on simulation and modelling in health care. J Simul. 3:130–140.

Bricker DA, Markert RJ. 2010. Night float teaching and learning: Perceptions of residents and faculty. J Grad Med Education. 2(2):236–241.

College of Nurses of Ontario. 2018. Practice guideline: RN and RPN practice: The client, the nurse, and the environment. Toronto (ON).

de Bruin AM, van Rossum AC, Visser MC, Koole GM. 2007. Modeling the emergency cardiac in-patient flow: An application of queueing theory. Health Care Manag Sci. 10(2):125–137.

Dobson RT, Stevenson K, Busch A, Scott DJ, Henry C, Wall PA. 2009. A quality improvement activity to promote interprofessional collaboration among health professions students. Am J Pharm Educ. 73(4).

England W, Roberts SD. 1978. Applications of computer simulation in health care. In: Highland HJ, Nielsen NR, Hull LG, editors. WSC 1978. Proceedings of the 1978 Winter Simulation Conference – Volume 2; 1978 Dec 4–6; Miami Beach, FL. Los Alamitos (CA); IEEE. p. 665–677.

Fone D, Hollinghurst S, Temple M, Round A, Lester N, Weightman A, Roberts K, Coyle E, Bevan G, Palmer S. 2003. Systematic review of the use and value of computer simulation modelling in population health and health care delivery. J Public Health Med. 25(4):325–335.

Günal MM, Pidd M. 2010. Discrete event simulation for performance modelling in health care: a review of the literature. J Simul. 4:42–51.

Hamrock E, Paige K, Parks J, Scheulen J, Levin S. 2013. Discrete event simulation for healthcare organizations: A tool for decision making. J Healthc Manag. 58(2):110–125.

Hu X, Barnes S, Golden B. 2018. Applying queueing theory to the study of emergency department operations: A survey and a discussion of comparable simulation studies. International Transactions Operational Res. 25:7–49.

Jacobson SH, Hall SN, Swisher JR. 2006. Discrete-event simulation of health care systems. In: Price CC, editor. International Series in Operations Research and Management Science Volume 91. New York City (NY): Springer; p. 212–252. 206:273–309.

Jun J, Jacobson J, Swisher J. 1999. Application of discrete-event simulation in health care clinics: A survey. J Operational Res Soc. 50(2):109–123.

Klein MG, Reinhardt G. 2012. Emergency department patient flow simulations using spreadsheets. Simul Healthc. 7(1):40–47.

Law AM. 2007. Simulation modelling and analysis. Boston: McGraw-Hill.

Lehaney B, Hlupic V. 1995. Simulation modelling for resource allocation and planning in the health sector. Perspect Public Health. 115(6):382–385.

Mandelbaum A, Momčilović P, Tseytlin Y. 2012. On fair routing from emergency departments to hospital wards: QED queues with heterogeneous servers. Manag Sci. 58(7):1273–1291

Powell ES, Khare RK, Venkatesh AK, Van Roo BD, Adams JG, Reinhardt G. 2012. The relationship between inpatient discharge timing and emergency department boarding. J Emerg Med. 42(2):186–196.

Python Software Foundation. 2017. https://www.python.org/.

Rasheed F, Lee YH, Kim SH, Park IC. 2012. Development of emergency department load relief area – gauging benefits in empirical terms. Simul Healthc. 7(2):343–352.

Rossetti MD, Trzcinski GF, Syverud SA. 1999. Emergency department simulation and determination of optimal attending physician staffing schedules. In: Farrington PA, Nembhard HB, Sturrock DT, Evans GW, editors. WSC 1999. Proceedings of the 1999 Winter Simulation Conference; 1999 Dec 5–8; Phoenix, AZ. Phoenix (AZ): IEEE. p. 1532–1540.

Sarno D, Nenni ME. 2016. Daily nurse requirements planning based on simulation of patient flows. Flex Serv Manuf J. 28(3):526–549.

Shenoy ES, Lee H, Ryan EE, Hou T, Walensky RP, Ware W, Hooper DC. 2018. A discrete event simulation model of patient flow in a general hospital incorporating infection control policy for Methicillin-Resistant *Staphylococcus aureus* (MRSA) and Vancomycin-Resistant *Enterococcus* (VRE). Med Decis Mak. 38(2):246–261.

Siddiqui S, Morse E, Levin S. 2017. Evaluating nurse staffing levels in perianesthesia care units using discrete event simulation. IISE Transactions Healthc Syst. 7(4):215–223.

Sundaramoorthi D, Chen VCP, Rosenberger JM, Kim SB, Buckley-Behan DF. 2010. A data-integrated simulation-based optimization for assigning nurses to patient admissions. Health Care Manag Sci. 13(3):210–221.

Wilson JCT. Implementation of computer simulation projects in health care. J Opl Res Soc. 32(9):825–832.