Using Poisson Modeling and Queuing Theory to Optimize Staffing and Decrease Patient Wait Time in the Emergency Department

Geralda Xavier¹,², Joseph Crane²,³, Michele Follen¹,²*, Wendy Wilcox¹, Steven Pulitzer¹, Charles Noon²

¹Kings County Hospital, Brooklyn, NY, USA
²Haslam College of Business, University of Tennessee, Knoxville, TN, USA
³Teamhealth Inc., Knoxville, TN, USA
Email: *Michele.Follen@mychhc.org


Received: January 13, 2018
Accepted: September 27, 2018
Published: September 30, 2018

Abstract

Introduction: Studies have shown Emergency Department (ED) crowding contributes to reduced quality of patient care, delays in starting treatments, and increased number of patients leaving without being seen. This analysis shows how to theoretically and optimally align staffing to demand. Methods: The ED value stream was identified and mapped. Patients were stratified into three resource-driven care flow cells based on the severity indices. Time observations were conducted for each of the key care team members and the manual cycle times and service rate were calculated and stratified by severity indices. Using X32 Healthcare’s Online Staffing Optimization (OSO) tool, staffing inefficiencies were identified and an optimal schedule was created for each provider group. Results: Lower Severity Indices (higher acuity patient) led to longer times for providers, nurses, patient care assistants, and clerks. The patient length of stay varied from under one hour to over five hours. The flow of patients varied considerably over the 24 hours’ period but was similar by day of the week. Using flow data, we showed that we needed more nurses, more care team members during peak times of patient flow. Eight hour shifts would allow better flexibility. We showed that the additional salary hours added to the budget would be made up for by increased revenue recognized by decreasing the number of patients who leave without being seen. Conclusion: If implemented, these changes will improve ED flow by using lean tools and principles, ultimately leading to timeliness of care, reduced waits, and improved patient experience.

Keywords

Poisson Modeling, Queuing Theory, Reduced Waits, Improved Patient Experience
1. Introduction

The ED is the hospital’s “front door” and is the primary source of inpatient admissions. EDs across the country are facing increasing challenges because of growing patient volume and decreasing hospital resources. Consequently, most EDs are overcrowded and unable to flex staffing capacity to meet demand.

The Institute of Medicine report defines several domains for quality of care in the Emergency Department, including: safety, patient centeredness, timeliness, efficiency, effectiveness, and equity [1] [2] [3] [4]. ED visits are going up annually and are often the safety net for medically underserved patients [5].

The flow from home to the ED is impacted by many factors that can be characterized in separate areas. “Flow In” is impacted by ambulance deployment, ambulance diversion, and appropriate referral/use of the ED. “Patient flow within the ED” is impacted by: the effectiveness of triage, the use of a “fast track” for low acuity patients, ED staffing, ED scheduling, available Intensive Care Unit and other inpatient beds, and the presence of an observation unit [6]. “Patient Flow out” is affected by family’s ability to take the patient home or by the skilled nursing facilities availability.

Long wait times, previously called crowding in the literature, have been shown to result in poor care for patients with severe pain and in delays in antibiotics administered for pneumonia [7] [8] (W. E. Fee C n.d.). Further, a review by Bernstein et al. of available literature showed that overcrowding is associated with increased mortality [9].

Not only are patients treated more slowly than desired, but after long waits, they leave or elope. These patients who leave without being seen are tragic problems for EDs. Baker et al. conducted a follow-up study of patients who left without being seen and found that 46% of those who left, actually needed immediate care, 29% needed care in the next 24 - 48 hours, and 11% overall were hospitalized within one week of leaving [10].

Part of the difficulty with studying crowding was the lack of consistent terms and definitions. Coalitions of researchers have convened and Welch et al. has published two dictionaries that precisely define the terms under study in 2011 and in updates in 2014 [11] [12] [13]. Now that measures are defined and benchmarks defined, EDs can begin measuring themselves against national standards.

In this study, we examine: the map of patients in the ED, the minutes spent by staff at each stage in the ED across the severity levels, patient length of stay, manual cycle times for each care team member stratified by patient severity index, the hourly service rate for each care team member, and examine our ED staffing of nurses, advanced practice providers, and physicians. The goal of this analysis is to optimally align staffing to the demand model for these key servers of care. If implemented, it is anticipated these models will improve ED flow, ultimately leading to improved timeliness of care, reduced waits, and improved patient experience.
2. Methods

2.1. Sample Sizes

There are two sample sizes for this work. For the observations of flow by severity index, a month of data from 2016, was randomly chosen using computer simulation. Approximately 150,000 patients are seen annually at Kings County Hospital. A subsequent analysis (data not shown) demonstrated that there is relatively little month to month variation. However, since the database and data analysis programs were robust, a random month was chosen and the data from a spreadsheet file and was entered into the Online Staffing Optimization (OSO) software program of X32 Healthcare.

Subsequently, time measurements of 30 patients (10 for each of three severity index clusters) were followed on randomly chosen weekdays and randomly chosen shifts for the analysis of minutes spent by patients in each area. These studies were planned by Xavier and Crane and were designed to focus on a single patient as they experienced the activities in the process. Different days and different shifts were sampled to avoid bias in the sampling. Three researchers timed the patients, Xavier and two Emergency Room residents, and care was taken to use the same methodology and to sample from all three shifts and each of the seven days (Data not shown).

2.2. Determining the Emergency Department’s Value Stream and Mapping It

In most hospitals, patients are triaged based on the Emergency Severity Index (ESI), a five-level emergency department triage algorithm that provides clinically relevant stratification of patients into five groups from 1 (most urgent) to 5 (least urgent) based on acuity and resource needs [14].

Based on this system, we developed three resource-driven care flow cells (patients stratified by acuity and resource intensity) were created, thus allowing patients to be streamed to a resource driven flow cell [15]. Each of these flow cells have dedicated staff in a specific geographic area within the ED. This process of dividing patients into interconnected flow cells has reduced wait times and decreased motion for patients compared to a model that does not split patients into acuity stratifications [14].

The high resource-driven flow cell involves patients triaged ESI 1 & 2. These patients are the most resource intense and are also allocated to a dedicated clinical team and geographic location. The medium resource-driven flow cell involves patients triaged as ESI 3. This flow cell is also staffed with a dedicated clinical team, including attending physicians, both emergency/internal medicine residents, nurses, a patient care technician, a clerk, and care managers to ensure this cell’s needs are met in a timely manner. The low resource-driven flow cell involves patients triaged as ESI 4&5. These low resource patients are typically “see and treat” and are allocated to a dedicated clinical team who only sees that type of patients. The team of providers for this flow cell includes a physician, an
advanced practice provider, a nurse, a patient care assistant, and a clerk. This flow cell allows relatively “well” patients to be seen in a timely manner (Figure 1).

Figure 2 depicts the complex environment of patient flow through the health care continuum. Bottlenecks may occur: 1) within an area (internally within the ED itself), 2) between areas (internally within the hospital such as finding beds for admitted patients to either ICU’s or medical/surgical beds), and 3) with outside facilities such as Skilled Nursing Facilities (SNF) or discharge to home. Figure 3 depicts patient’s arrival into these treatment flow cells used for this study.

2.3. Time Observation of Staff

The authors observed care in the ED when they were not otherwise on duty. They defined flow cells and commenced with time observation studies. These were conducted by the first author and two research assistants who observed each of the key care team members: clerks, patient care assistants (PCA, equivalent to nursing assistants), nurses, and providers (including both Advanced Practice Providers and Physician Providers). Data were collected over all three shifts and over different days of the week. This time data was collected as de-identified data and this study was considered Internal Review Board exempt. Time observations (also called time studies) are important for accurately measuring the work load of each resource-driven flow cell. A time observation captures all the provider’s steps in the current state and the cycle time for each task separately. The procedures are outlined in the Crane text [16].

After randomization of the day of the week and shift, the three researchers...
Figure 2. Patient flow across care the continuum in the Emergency Department (ED) detailing flow to: The Observation (Obs) unit, the Inpatient unit (Inpt), the Care Management Unit (Care mgmt.), and eventually to either Home or to a Skilled Nursing Facility (SNF).

Figure 3. Treatment Flow Cells (FC) showing how patients arrive and then are triaged based on the emergency severity index and resources needed for their care before exiting to the departure flow cell.
followed patients carefully through all the steps in the ED. Ten patients were observed for each of the three severity index clusters. Data was collected on all the key care team members. The following materials and equipment were needed: stop watch, clip boards, a time-observation sheet, paper, and pencil.

Rather than resetting the stop watch to zero after each step, the time keeper announced the total elapsed time as the care team member performed each step. We calculated the individual time for each step excluding the wait times. The time recorder wrote the time for each process element on the time observation sheet. Data from the observation sheet was entered into a spreadsheet.

2.4. Service Rate, Wait Time, Queuing Theory

The Arrival Rate is the mean number of arrivals per unit time (usually per hour or per day). The Manual Cycle Time is equal to the time to complete the process minus any wasted time such as wait time. The Service Rate is the mean number of customers that can be served at 100% utilization by each individual server per unit time (usually per hour or per day). After the time observations were completed, the cycle times were calculated for each key team member. The manual cycle times for each key care member are shown in Tables 1-3.

With the Manual Cycle Times known, the service rate can be calculated for each team member. A server’s service rate is typically calculated based on an hourly period by dividing 60 minutes by the MCT. Therefore, for the medium resource flow cell, the physician service rate is 1.83 patients per hour. In other words, if eight patients arrive per hour with three physicians on duty, only five patients can be seen per hour, meaning three patients are waiting every hour because the rate of arrival (eight patients per hour) exceeds the service rate (five patients per hour). Thus, for every hour the system runs this way, the queue length will grow an average of 3 patients; and assuming a steady state of arrival, about 72 patients will wait in a 24-hour period.

2.5. Queuing Analysis

Queuing is a mathematical method of analyzing congestions and delays of waiting in lines or queues. Queuing theory is based on Poisson modeling. A thorough mathematic description can be found in this reference [17]. Ideally, queuing theory allows examination of every component of waiting to be served, including the arrival rate, service rate, number of servers, number of system places, and number of patients.

Queues are found nearly everywhere in healthcare. In the ED, we see queues on patient arrival at check-in; queues for triage; queues for providers; and queues for admission and discharge. For example, the first servers who patients encounter on entering an ED are a registration clerk followed by a nurse. When arrival rates are greater than the servers’ capacity or there is considerable variation in arrival patterns and/or service times, queuing occurs. This happens because the servers are still caring for patients who arrived earlier thus, the newly arriving patients must wait in a line.
Table 1. Data collected in the high severity index unit from 10 patients (measured in patient minutes spent in each area) and averaged. Patient Care Assistants are nursing personnel at this institution.

<table>
<thead>
<tr>
<th>ESI 1/2 Cell</th>
<th>People</th>
<th>Provider</th>
<th>Reg. Nurse</th>
<th>Patient Care Assistant</th>
<th>Clerk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Results Review</td>
<td>2.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>History &amp; Physical Exam</td>
<td><strong>10.0</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consults</td>
<td>3.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disposition</td>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedures</td>
<td>10.0</td>
<td><strong>2.0</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Place Orders</td>
<td>2.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assessment/Reassessment</td>
<td>2.0</td>
<td>2.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Documentation</td>
<td>2.5</td>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient &amp; Family Communication</td>
<td>2.5</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine Administration</td>
<td>7.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labs/Bloodwork</td>
<td>3.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vital Sign Measurement</td>
<td></td>
<td></td>
<td>2.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EKG</td>
<td></td>
<td></td>
<td>4.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Point of Care/Urine Collection</td>
<td></td>
<td></td>
<td>2.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Registration</td>
<td></td>
<td></td>
<td></td>
<td>4.7</td>
<td></td>
</tr>
<tr>
<td>Patient Comfort Care</td>
<td>2.9</td>
<td>2.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient Transport</td>
<td>2.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Triage</td>
<td></td>
<td></td>
<td>5.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Manual Cycle Time Total: 35.9 25.7 14.3 4.7

2.6. Staffing to Demand

To meet the facility’s goals as described in the True North Metrics [1], X32 Healthcare’s Proprietary Online Staffing Optimization (OSO) tool was used. [18] The OSO tool is based on the use of Poisson Modeling and uses emergency department patient metrics, staffing inputs and volume information, and common emergency department statistical data to highlight demand versus capacity and hourly patient arrival distributions. The OSO tool can easily compare current and future staffing profiles, thus allowing teams to effectively allocate resources to reduce bottlenecks and better meet patient demand.

Theoretically these models can be used to understand the demands made of clerks, patient care assistants, nurses, advanced practice providers, and physicians by both hour of the day and day of the week. This information can then be compared with the current staffing levels to identify inefficiencies and create an optimal schedule for each service group. Our ED sees 140,000 patients annually.
Table 2. Data collected in the medium severity index unit from 10 patients (measured in patient minutes spent in each area) and averaged. Patient Care Assistants are nursing personnel at this institution.

<table>
<thead>
<tr>
<th>ESI 3</th>
<th>People</th>
<th>Provider</th>
<th>Reg. Nurse</th>
<th>Clerk</th>
</tr>
</thead>
<tbody>
<tr>
<td>History &amp; Physical Exam</td>
<td>10.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Precepting/Case Presentation</td>
<td>5.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedures</td>
<td>5.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Place Orders</td>
<td>1.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consults</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Review Results</td>
<td>2.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disposition Note</td>
<td>4.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disposition</td>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medication</td>
<td>4.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vital Signs</td>
<td>2.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV Placement</td>
<td>5.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient Transport</td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assessment/Reassessment</td>
<td>2.0</td>
<td>1.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient Gowning</td>
<td></td>
<td>1.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Registration</td>
<td>4.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Point of Care</td>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrocardiogram</td>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bloodwork</td>
<td>3.8</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Manual Cycle Time Total: 32.8 24.5

We randomly chose a month of data between October of 2016 and October of 2017 for part of the analysis and a random week of data for the staffing analysis.

2.7. Cost Analysis

Cost analysis is critical to operations analysis. To recognize the necessary improvements, cost must be considered. In this analysis, we based calculating the reimbursement lost on patients who left without being seen and compared that with the increased costs of staffing flexibly to meet the actual demand. These calculations were made on widely acknowledged salaries for various figures on the patient care team: clerks, patient care assistants, nurses, advanced practice providers, and physicians. The hourly rates pertain to urban areas in the Northeast.

3. Results

3.1. Timed Observation of Group

The roles of each care team member were identified and mapped before the time
observation data was collected for ten patient encounters in each flow cell over randomly selected days and all three shifts. The aggregated data was computed and is shown in Tables 1-3, stratified by the Emergency Severity Index. As expected, the time spent in each area almost always increases as ESI decreases (i.e. as acuity increases).

### 3.2. Manual Cycle Times, Service Rate, and Queues

We calculated manual cycle times and service rate for each of the three severity index cells for clerk, patient care assistants, nurses, advanced practice providers, and physicians. Examining the data in this way allows one to see how many staff of each type are needed per hour to keep up with the demand. Table 4 shows the manual cycle time and hourly service rate by each care team member stratified by the severity index of the flow cell.

### 3.3. Length of Stay

The flow changes over each hour and as patients remain in the ED, they remain
in a bed that cannot be used for another patient. To determine the queuing effect on patients arriving at our busy ED by hour of the day and day of the week, data was entered into the Queuing Flow Simulator developed by Ms. Clay and Dr. Noon. Figure 4 shows that the length of stay through the physician queue, as estimated by discrete event simulation, varies throughout the day and ranges from less than one hour to over five hours for some patients.

3.4. Using the Flow Data and Hourly Service Rate to Achieve Staffing to Demand

In Figure 5, the red bars represent patients in the high and medium resource-driven cells, and the green bars represent patients in the low-acuity resource-driven

Table 4. Data collected on 30 patients showing the Manual Cycle Time and hourly Service Rate for each member of the care team stratified by the Emergency Severity Indices.

<table>
<thead>
<tr>
<th>Resource Flow Cell</th>
<th>Care Team Member</th>
<th>Manual Cycle Time in Minutes</th>
<th>Hourly Service Rate in Patients/Hour</th>
</tr>
</thead>
<tbody>
<tr>
<td>High (ESI 4 &amp; 5)</td>
<td>Provider</td>
<td>35.9</td>
<td>1.67</td>
</tr>
<tr>
<td></td>
<td>Nurse</td>
<td>25.7</td>
<td>2.33</td>
</tr>
<tr>
<td></td>
<td>Patent Care Assistant</td>
<td>14.3</td>
<td>4.20</td>
</tr>
<tr>
<td></td>
<td>Clerk</td>
<td>4.7</td>
<td>12.77</td>
</tr>
<tr>
<td>Medium (ESI 3)</td>
<td>Provider</td>
<td>32.8</td>
<td>1.83</td>
</tr>
<tr>
<td></td>
<td>Nurse</td>
<td>24.5</td>
<td>2.45</td>
</tr>
<tr>
<td></td>
<td>Clerk</td>
<td>4.7</td>
<td>12.77</td>
</tr>
<tr>
<td>Low (ESI 1 &amp; 2)</td>
<td>Provider</td>
<td>19.4</td>
<td>3.09</td>
</tr>
<tr>
<td></td>
<td>Nurse</td>
<td>12.4</td>
<td>4.84</td>
</tr>
<tr>
<td></td>
<td>Patient Care Assistant</td>
<td>7.3</td>
<td>8.22</td>
</tr>
<tr>
<td></td>
<td>Clerk</td>
<td>6.9</td>
<td>8.70</td>
</tr>
</tbody>
</table>

Figure 4. Length of stay analyzed by hour of the day of patients seen in the Emergency Department using data from 2016 from over 140,000 visits annually (110,000 adult visits and 30,000 pediatric visits annually).
cells. The lower acuity demand is highest Monday through Thursday with peak arrival between 9 a.m. and 12 p.m. and lowest arrival between 11 p.m. and 6 a.m. Except for noted lower volumes, the high acuity and low acuity demands show a similar pattern of arrival. The patient demand will translate into provider workload for each given segment, which is discussed in more detail in the section in the following staffing analysis.

Table 5 shows the current patient metrics used in the model. These values are averages from a month of data chosen randomly between October 2016 to October 2017 provided by our internal scheduling software. Although 2016 and 2017 data was selected for this analysis, the data for the previous five years is based on 150,000 patient visits annually.

The percentile value in the third column of Table 5 compares these metrics with the EDs in the Centers for Medicare and Medicaid Services’ (CMS) data-
base. For example, consider the length of stay (LOS) discharged value of 255. The percentile score of 1.51% indicates that 98.49% (100% - 1.51%) of emergency departments in the CMS database have shorter length of stay. Our observations suggest that our patients have long length of stays and wait times. Furthermore, the increased length of stay requires more nursing time.

3.5. Nurse Staffing

**Figure 6(a)** shows the direct care nurse-resource demand based on the data in

---

**Figure 6.** (a) Current Nurse Demand in the Emergency Department hourly over 24 hours by day of the week with Sunday on the far left and Saturday on the far right; (b) Recommended staffing based on the hourly flow of patients over 24 hours by day of the week with Sunday on the far left and Saturday on the far right.
the input module in OSO. The y-axis displays total demand (in number of nurses needed), and the x-axis is the hour of the day and day of the week. The red line profile represents the current nurse staffing based on the scheduled nurse assignments, using 0.9 patients per hour as the nurses’ service rate. The above analysis reveals the following: 1) the demand for nursing resources far exceeds capacity during the week days; 2) on weekends, the demand almost matches the current staffing profile; and 3) there is excess capacity on the night shift.

The new analysis created by our model is shown in Figure 6(b); ideal staffing would be represented by the blue line. The new schedule adds nurse capacity during the days to fully meet the demand. It also removes hours from the overnight periods and allocates hours to necessary time segments. To achieve this new schedule, an additional 225 weekly hours would have added to this model.

3.6. Advanced Practice Provider Staffing

Figure 7(a) shows the advanced practice provider resource demand. These providers primarily work in the low acuity unit so that workload is governed by the percent of ESI 4 and 5 inputs. The y-axis displays total demand (in number of APPs needed), and the x-axis is the hour of the day and day of the week. The red line profile represents the current APP staffing levels. Based on the current schedule, which was entered into the OSO staffing model, we used the APP’s current productivity of one patient per hour as the service rate for the low-resource flow cell.

The above analysis reveals the following: 1) a spike in unmet demand from 11 a.m. to 3 p.m. on most days; 2) a mismatch in demand and capacity, leading to excess capacity on the front end and resulting in waste; and 3) a significant amount of unmet demand on the back end possibly leading to patients’ leaving without being seen due to long waits. Figure 7(b) shows the ideal schedule.

When this schedule was created, more consideration was given to shift lengths and start times. Shift lengths are 10 or 12 hours because of contractual constraints, making it difficult to truly optimize staffing to meet the demand. However, the new schedule has some excess capacity that could be used for patient activities such as positive culture follow-ups and post-discharge follow-up calls. A better-fitting scheduling profile could be obtained by allowing these providers to work eight-hour shifts.

3.7. Physician Staffing

Figure 8(a) shows the physician workload, which is governed by the percent of ESI 1, 2, and 3 inputs. Furthermore, any low acuity advance practice provider leftover demand not covered rolls over to physicians and is included in the physician-demand chart. The y-axis displays total demand (in number of physicians needed), and the x-axis is the hour of the day and day of the week. The red line profile represents the current physician staffing levels based on the actual schedule and is then entered the OSO staffing model. The model uses 2.4 patients per
hour as the physicians’ high resource flow cell service rate and 2.7 patients per hour as the medium resource flow cell service rate.

This analysis shows the following: 1) the provider resource in staffing is not aligned with demand with too much capacity on the front end; 2) the demand for physicians far exceeds capacity almost daily, likely leading to increased number of patients who leave without being seen as well as long waits to see a provider; and 3) excess capacity is noted during night shifts. The new proposed schedule better aligns physician resources to meet demand (see Figure 8(b)).
3.8. Cost Analysis

In the staffing-to-demand analysis, significant gaps in staffing for advanced practice provider and nursing were observed. To align patient demand with adequate nursing resources, 225 additional hours would be added when creating the recommended profile; 205 hours were added to APP role and 15 hours were added to physician hours. The resulting costs can be viewed as salary/hour of additional time minus increased revenue from having less or no patients leaving without being seen.

For example, a nurse’s average pay rate of $47/hour in New York State and is
equivalent to $10,575 increase in cost. An advanced practice provider pay rate of $70/hour is equivalent to 205 * 70 = $14,350 increase in cost. A physician’s pay rate of $200/hour is equivalent to 15 * 200 = $3000 increase in cost. Considering these changes ($10,575 + $14,350 + $3000), the new weekly cost is $27,925 more than the current model, almost $1.5 million annually.

In the current model, the APP service rate is one patient per hour, accounting for the need to increase hours to meet demand; however, if we benchmarked APP at two patients per hour service rate, we would decrease the weekly hours by 285, equivalent to a 285 * $70 = $19,950 decrease in weekly cost.

Let us assume that current rate of patients who leave without being seen is about 8%. If we assume each discharged patient represents an average charge of $250, an additional revenue of $2.8 M can be generated with this new staffing-to-demand model, thus increasing the ED’s profit margin.

4. Discussion

Problems encountered in the ED affect the entire hospital because the ED is the gateway to trauma surgery, intensive care unit admission, inpatient admission, and the observation unit. The ED also impacts the community with the need to transfer treated patients to their homes or to skilled nursing facilities. Systematic reviews by Hoot, Wiler, and Oredsson examine factors that contribute to crowding, the effects of crowding, and the solutions to crowding [19] [20] [21]. These problems are not unique to the US but exist worldwide. We have confined the discussion to the US, but investigators in the United Kingdom, Canada, and Australia have produced equal amounts of scholarly work.

From these reviews have come three interventions which have been widely adopted: rapid triage, fast-track, and observation units. White et al. conducted a study measuring pre- and post-intervention and showed that rapid triage and fast track decreased length of stay for both admitted and released patients and decreased those who left without being seen [22]. Chan et al. conducted a similar study of both interventions and show significant improvements in length of stay and patients leaving without being seen [23]. Rapid triage, fast-track, and observation units are already implemented in our ED.

Rapid Triage and Fast Track have decreased those who leave without being seen. Hsia et al. studied 9.2 million ED visits to 262 hospitals in California and found that the percentage of those who left without being seen ranged from 0 to 20% and average 2.6%. Factors that increased the number were poorly insured patients, county hospital ownership, having a trauma center, and having a teaching affiliation [24]. Baker et al. reported on the follow-up study of those who left without being seen showed that they left after an average of 6 hours of waiting [10].

Queuing theory addresses patient arrival and staffing and has been reported to be helpful in several studies. Like our use of Queuing to improve staffing, Green et al. showed that despite increasing numbers of arrivals, the provider staffing
changes resulted in 22% less patients leaving without being seen [25]. Our study suggests important staffing changes could result in similar findings.

Increasingly the Centers for Medicaid and Medicare and the Agency for Healthcare Research and Quality have worked with the Joint Commission to suggest reasonable treat and release times. They have focused on four hours as a reasonable time limit since 2012. They are routinely surveying patients about their experiences in the emergency department [15] [26] [27].

Fee et al. reviewed 2008 National Hospital Ambulatory Medical Care Survey data and stratified the data by safety-net versus non-safety-net hospitals. This sample included 33,134 patient records and 396 hospitals in the US. The median length of stay for patients admitted for safety net hospitals was 269 minutes and for non-safety net 281 minutes. The median length of stay for patients discharged for safety net hospitals was 156 minutes and for non-safety net 148 minutes. The median length of stay for patients in observation for safety net hospitals was 355 minutes and for non-safety net 298 minutes. Finally, the median length of stay for patients transferred from safety net hospitals was 235 minutes and from non-safety net 239 minutes. The data was from 2008, before the widespread use of rapid triage and fast track. Most of those numbers are above four hours [26].

If this new staffing model is implemented, the benefits will be far reaching in terms of the following: 1) improved timeliness of care, translating to reduced waits; 2) improved quality by getting the right team to patients sooner, resulting in fewer patients leaving without being seen; 3) improve revenue capture by decreasing the number of patients leaving without being seen; and 4) meet the Centers of Medicare’s treat and release times. This would be increased Full Time Equivalent staffing at the physician and advanced practice provider level but in 8 hour rather than 12 hour shifts to keep this adjustment cost-effective.

Acknowledgements

We thank the patients who use the Emergency Department at Kings County and hope these studies help them. We also thank Sheldon McLeod, our CEO; Opal Sinclair-Chung, our CNO; Graham Gulian, our COO, and Steve Pulizter, our CMO, for having the vision to support continual change for betterment in our system.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Institute of Medicine (2006) Hospital-Based Emergency Care: At the Breaking Point. Institute of Medicine, Washington DC.


https://doi.org/10.1016/j.annemergmed.2009.05.021


