

Measuring Medical Provider Strain: Explaining Intensive Care Workload under Highly Variable Demand

Todd Link, Michael Rieker, Ryan Silva, Kate Zimmerman, Johnathan Dulin, Gregory Steeger, and Bradley Warner

Operations Research Program, United States Air Force Academy, Colorado, USA

Corresponding author's Email: C18Todd.Link@usafa.edu

Author Note: Todd Link, Michael Rieker, Ryan Silva, and Kate Zimmerman are senior cadets at the US Air Force Academy, collectively working on this project as part of a year-long operations research capstone course. The student authors would like to extend thanks to the advisors involved with this project as well as the client organization, Children's Hospital of Colorado.

Abstract: The Pediatric Intensive Care Unit at the Children's Hospital of Colorado is the premier provider of critical care for young patients with acute illnesses in a five-state region, but is overworking providers. Given the variable demand and the unpredictability of an intensive care environment, the Pediatric Intensive Care Units struggle to predict how much strain the unit will experience. This study develops a composite measure of strain felt by the medical staff of the Pediatric Intensive Care Unit to afford insights to providers. Discussions with the Pediatric Intensive Care Unit Team led to the generation of a strain "stoplight", to discriminate between high, medium, and low strain days. The strain metric serves to identify the days that present greater than normal strain. Through ordered logistic regression, the analysis finds that admissions and census data contribute most to the unit's strain, with admissions contributing twice as much to strain per unit.

Keywords: Optimization, Scheduling, Strain, Patient Flow, Hospital Operations, Resource Allocation

1. Introduction

Personnel scheduling is an area in the medical field which has experienced improvements through optimization. For example, nursing operations have been studied many times and with analytically created schedules (Kellogg and Walczak 2007). However, this research has not yet successfully been applied to Pediatric Intensive Care Units (PICU), largely due to the highly variable patient demand that these branches work under. This demand, combined with other factors addressed in this study, creates a strain on providers which has not been defined or standardized within the healthcare industry; the lack of agreement (Soltani et al 2015) makes further analysis or recommendations for operational improvements difficult.

One such PICU is attached to the Colorado Children's Hospital (CHCO) in Aurora, Colorado. Currently the CHCO PICU is one of the premiere children's and teaching hospitals, and thus must service a heavy demand and a wide variety of patients. The PICU is focused on providing the best possible care for each of its patients, and it helps children and families recover from serious illness or injury. The PICU serves an eight-state region around Colorado, and ultimately aims to save the lives of suffering children. In developing a new definition of strain on the providers, this project will attempt to aid in the mission of the CHCO PICU and alleviate pressure on providers.

1.1 Current State

The PICU can service 32-48 patients depending on its seasonal configuration. Providers, including doctors and advanced practitioners, are assigned to three different teams: red, white, and blue. The red and white teams include six providers, while the blue team includes only two. This smaller team is used as a flexible entity to interface with patients and other providers as needed in the summer. In the winter, the blue team increases to a six provider team like the red and white teams. Providers include attending physicians, who are considered pediatric experts; fellows, who are training to become physicians; residents; nurse practitioners; and physician's assistants. The residents, nurse practitioners, and physician's assistants are frontline providers, who perform most of the charting, researching, and presenting the patient complications. PICU patients include severely ill or injured children from an eight-state region around the Rocky Mountains. About 3,000 patients come through the PICU each year. These patients require a wide range of treatments and constant treatment adjustments.

Current scheduling methods do not account for any factors other than the number of beds filled (census). The number and type of base providers hired and assigned to teams remains the same throughout the year. The only adjustment used to account for increased seasonal demand involves the hiring of additional providers seasonally. During the winter, the PICU hires additional providers due to increased demand. Therefore, rapid staffing changes must be made in order to accommodate for growth.

1.2 Problem Statement

In this project, we attempt to answer the following question: how can we best quantify the strain that highly-variable patient demand in the PICU places on the providers? Our strain metric measures the workload, meaning the workload that the medical staff deals with on a daily basis in order to effectively handle patients and operations in the PICU.

The strain metric quantifies the level of burden the providers (at each level) face. Many different factors potentially go into this variable including number of patients, level of patient illness, number of orders placed on patients' charts, and/or number of phone calls within the unit. It is also important to note that the strain felt is distinct for an attending physician than for an Advanced Practice Provider (APP) or resident physician. The establishment of this strain method will facilitate the creation of a scheduler in the future for providers which will ensure the PICU is best able to meet the needs of their patients and providers.

1.3 Related Work

Analyses of the healthcare industry have focused primarily on the inherently variable patient demand. Industry standards come directly from studies which model patient flow and assess the hospital response to demand. Utilizing these studies not only offers insight into the factors affecting strain, but serves as a framework for modelling the daily operations of medical care.

Tunc (2014) discusses how the decision making process that doctors use, which involves precedent and intuition, often lacks consistency due to a lack of analytical backing to the medical decisions that are made regarding patients. Providers within the American medical system have raised concerns that the ad hoc medical decision making process may constitute a source of strain, but analysts predict that a more consistent approach regarding patient decision making will reduce strain. Further, a more disciplined and analytical process to assigning patients to teams in the PICU will produce a positive effect, as demonstrated by Rais and Viana (2011) in their study of healthcare seasonal demand forecasting. Rais and Viana's work should be helpful because they examine how the time of year in conjunction with demographics affects the level of demand for medical care. However, Kellogg and Walczak (2007) present explanations for the disparity between academic models and practical models including reasons such as lack of engagement with the client and an incorrect understanding of hospital operations; it is important to keep the stakeholders involved throughout the research in order to ensure a useful end-product.

An important component of this project involves measuring strain as this comprises a major piece of the analysis requested by the PICU. In related research, few studies attempt to create any such metric, or to elaborate on their metrics used to determine demand. Soltani et al (2015) define strain as "a temporally varying influence on a given ICU's ability to provide high-quality care for every patient who is or could become a patient in that ICU on a given day". The authors hold that the current solution to relieve strain by opening more beds is not the best for long term improvement of care. In addition, Soltani et al argue that a single measure of strain will not be enough to accurately describe the complexities of the metric.

The earliest attempt to quantify demand was the delay tool, created by Klein, Beshansky, and Selker (1990). The delay tool serves to classify any medically unjustified delays using several categories including scheduling, consultation, physician, and discharge. Faddy and McClean (1999) explain this strain metric through phase-type distributions. This study uses the phase-type as a metric for severity of illness, sorting patients into three categories: short-stay, medium-stay, and long-stay. Using this metric, Faddy and McClean were able to define a distribution for patient demand and time spent in the hospital. Our study will update this metric, and enhance the strain measures used for forecasting and prediction, introducing an optimized basis for prediction. More recently, a study by Kadri et al. (2013) focuses on developing a way to accurately represent strain situations that occur within hospitals. The authors hope to better understand the strain situations and use this information to improve emergency department operations and management. Instead of phase-type distributions, this study uses an architecture for an optimized and reactive control (ORCA) approach to model hospital system behaviors and situations defined by strain.

Considering this, few measures of strain that have been used in previous analyses are comprehensive or overly successful, as described by KL Teow (2009). Production of a more robust strain metric offers more complete scheduling practices in the future, and it will augment studies already completed.

Possibly the study most closely related to our efforts, Kolker (2008) focuses on being able to correctly model Intensive Care Unit (ICU) patient flow. The goal of his simulation model is to determine the connection between daily elective surgeries and the number of patients diverted from the ICU, either sent to other departments or other hospitals. A similar model of patient flow can represent the constantly changing patient demand of an intensive care unit.

2. Methodology

Developing a model for the PICU requires a multi-step process. Since the current metric does not do an excellent job at measuring the level of work, a better metric must be created. To do this, data was first separated into strain inputs (factors believed to contribute to the level of strain within the unit) and negative outcomes (adverse events potentially representing an undesirable result of high strain within the unit). A good strain metric would theoretically measure inputs into the PICU environment (i.e. census, admits, discharges, orders) and then determine negative outcomes that are a result of these inputs. Based on this information, Principle Components Analysis (PCA) examined how the features vary with regards to each other. PCA is a dimensionality reduction method which conserves as much of the variability in the data as possible while projecting onto a lower-dimensional space. In this analysis, PCA is used to provide insight on the relationship between the variables construed as inputs and those categorized as adverse events.

The team then conducted directed discussions with three medical teams in the PICU to ground our understanding of current operations. Using these directed discussions, which focused on strain as it relates to each different provider type, the team determined what variables have the greatest impact in creating a busy, or strained, day from the provider perspective. These discussions illuminated the number of Rapid Response Team (RRT) calls as a potential indicator of strain. The RRT variable records the number of times the PICU RRT is called by another ward in the hospital to evaluate a patient for potential admission into the PICU. One of the most challenging issues in this project is finding a reliable indicator of how much strain the PICU is actually experiencing each day. RRT calls answer this problem by indirectly approximating strain. The discussions with providers revealed that the rest of the hospital acts as a barometer of the PICU's strain. When the PICU is under a high workload, the hospital can sense its stress. Other units in the hospital react to this by holding patients that may need to go to the PICU within their units. Those patients often deteriorate, resulting in RRT calls for patient evaluation and potential PICU admission. Therefore, when the strain on the PICU is high, the number of RRT calls generally increases. Bad readmissions into the PICU and diverts away from the PICU also reflect the negative outcomes of strain. When the PICU is stretched to its capacity, providers are likely to transfer patients to other departments or send them home before they are ready. Bad readmissions occur when the patient is readmitted into the PICU within 24 hours, and the hospital determines that they should not have been released initially. There are likely to be more bad readmits when the strain is high and the providers feel pressure to open beds for potential new patients in severely critical status. Similarly, a divert signifies extremely high levels of PICU strain. A divert occurs when the PICU cannot admit another patient and thus must turn them away. Because of the nature of the PICU's mission (treating ill children) a diversion of a child represents extremely significant strain.

The providers also discussed the factors that drive whether a day is particularly high or low with regards to the strain placed on providers. Based on these discussions, number of admissions, orders, orders per patient, discharges, percent capacity, and the season all had impacts on the level of severity of strain. Any admission involves a high level of paperwork and takes providers time to familiarize themselves with a case. For this reason they represent a significant amount of work on the unit. Orders impact strain because more orders typically represents more work as the unit has to process more work with more orders. Orders per patient serves as a relatively good proxy for the acuity level of each patient with higher acuity patients representing more strain on the unit. Discharges also involve significant amounts of paperwork and put strain on the unit. Whenever the census numbers are close to capacity levels, there is significant added strain because doctors become concerned with moving patients out of the PICU in addition to the already strenuous job of treating the patients. Finally, according to the providers, the season (winter versus summer) impacts their operations; strain is universally regarded as higher in the winter.

Using this information, the team proceeded to create a metric to indicate strain currently being experienced by the unit. This metric is displayed as a 'stoplight' with three potential values: 0, 1, or 2. A value of 2 is representative of higher levels of strain, while a value of 0 is indicative of low levels of strain. The stoplight method of analyzing performance is standard in the medical field and the team determined it to be the best way of accurately showing the severity of the strain in an interpretable manner.

The team determined RRT calls to be the most clear indication of strain on the unit. For this reason, the foundation of the stoplight is this variable. After the RRT data is separated into high, medium, and low days based on resemblance to a poisson distribution, two variables were added to ensure a more accurate representation of high strain days. Any day with either a bad readmission or a divert documented was automatically placed in the 2 category for the stoplight. This is because

these two events invariably indicate a very high strain day for the PICU. The Poisson distribution and the two indicators of severe strain completed the stoplight variable.

This stoplight first needed to be verified to ensure it accurately reflected days that were high strain and low strain. First, providers at the PICU confirmed that this way of reflecting good days and bad days made sense. This high-level holistic check ensured that the metric made sense to those who deal with the strain every day. Next, K-means clustering separated the data into different clusters. The K-means clustering algorithm is a technique that relates all of the unlabeled observations to each other by Euclidean distance in a normalized n-dimensional vector space (where n is the number of features). Therefore, it is meant to find groups not explicitly labeled in the data. Similar observations are clustered together based on the variation within the data.

With these two confirmations on the strain stoplight, the team ran ordered logit regressions to determine which input variables determined by PICU staff had the largest effect on the stoplight. Ordered logit regression is a multinomial (the response variable is categorical but not binary) logistic regression model where the response is not only categorical, but ordered. In this specific instance, the response variable progresses from a 0 (low strain) to a 1 (medium strain) and then to a 2 (high strain). There is a proper order to the categories, as a 0 is distinct from a 2 in a different way than a 0 is distinct from a 1. The ordered logistic regression takes this into account.

$$Stoplight = \begin{cases} 2 & \text{if } RRT \geq 4 \cup Diverts \geq 1 \cup Readmits \geq 1 \\ 1 & \text{if } 2 \leq RRT < 4 \\ 0 & \text{if } RRT < 2 \end{cases} \quad (1)$$

With this final analysis, the team determined the relative weights of inputs into the strain variable and were able to provide insight to the PICU.

3. Results and Analysis

Following the methodology described in section 2, the team came to several conclusions and insights regarding the PICU environment and its relationship with strain.

First, we checked the accuracy of the current strain metric, census data. The use of this correlation is based primarily on advice from subject matter experts, who confirmed that as the unit experiences enhanced strain (patients requiring more attention, greater number of patients, more patient intakes, etc), the number of orders made by providers increases. The correlation between the number of orders and the number of beds is 0.687. This shows that higher number of beds does not always mean more work for the PICU, as patient acuity has a significant impact on unit strain, and confirms that this industry requires a more robust strain metric.

As described in the methodology section, the first method used to analyze the data involves unsupervised learning and Principal Components Analysis (PCA). This technique creates a visualization of the data by reducing the dimensionality while preserving as much of the variability as possible. Using PCA, we find that proxy measures for strain and adverse events are not directly dependent; the PCA reveals that the two categories of variables do not vary together. PCA demonstrated that the potential inputs for strain (number of orders, census, and admits/discharges) all vary together along one principal component, while outputs like medical errors, adverse events, bounce-backs, and deaths all vary together along principal component two. This analysis shows that the inputs and the outputs appear to be independent of each other, due to their principal component weightings being close to perpendicular.

The next component of our analysis creates a legitimate method of verifying strain levels for each day of the 3 years of data provided by the PICU. Although there is little variation between the the inputs and potential outputs, RRT represents a simple way to accurately determine if a day had significant strain. The RRT distribution is similar to a poisson distribution, with an initial increase in number of calls and then a rapid decrease.

Using poisson distribution percentiles, separations were made to designate high, medium, and low strain days. Using information from the providers, we determined that there should be approximately 30% high strain days. After this data was identified, we included the bad readmission and divert variables as discussed in section 2. This completed our strain stoplight metric and can be seen below.

Confirmation of the stoplight was made by the providers, as discussed above, and by using K-Means Clustering. The resulting cluster plot separates the data into 3 clusters, and the data points are colored based upon the value of the corresponding augmented RRT variable value (either 0, 1, or 2). The clustering algorithm uses normalized scores for orders per patient, number of admits, number discharged, census, and percent capacity in order to classify the observations. The result demonstrates that the components explain 88.1% of point variability.

This analysis groups 72% of the days labeled as a 2 (a bad strain day) in the same cluster. Providers are primarily concerned with these high strain days; this clustering demonstrates that the days labeled as high strain do in fact have similar input variables, which would lead to similar amounts of strain. These two methods confirmed the stoplight's relationship to strain, and allowed for continuation of analysis.

Using ordered logit regression, the relative weights of the input variables relative to stoplight were determined. The input variables used in this model were those identified by the providers as factors likely to impact how bad strain was on any given day.

$$y_i = \beta_0 + \beta_1 orders_i + \beta_2 orders\ per\ patient_i + \beta_3 admissions + \beta_4 discharges_i + \beta_5 census_i + \beta_6 percent\ capacity_i + \beta_7 winter_i + \mu$$
$$0 \leq i \leq 1095 \quad (2)$$

Where y_i represents the strain felt on day i .

The regression ultimately reveals that census and number of admissions are the most important factors in explaining the level of strain the PICU experiences on a given day, with admissions have a per-unit impact that was twice as large as census. That is, 1 additional admission was roughly equivalent to 2 additional beds filled in terms of the impact on the unit's strain.

4. Conclusions and Future Research

Utilizing variables obtained from day-to-day operations of the CHCO PICU, this analysis produced a stoplight metric. This metric offers an alternative to industry-accepted practices which associate strain with hospital capacity alone. Through analysis of the PICU data, the team determines that strain and capacity are not sufficiently related to create a predictive model or accept this metric as a standard for further research. Instead, the strain metric proposed by this study utilizes variables determined through both data analysis (including unsupervised learning and clustering) and PICU provider input.

Analysis of the strain metric compared with the data collected from 2015-2017 in the CHCO PICU suggests certain adverse events, such as RRT calls and diverts, in the unit are closely associated with the varying levels of strain, while other adverse events, such as deaths and respiratory arrests, are almost entirely independent. Consideration of these different levels of related adverse effects and associated levels of strain led to the creation of a stoplight measure, which uses different variable inputs that lead to high, medium, or low levels of strain.

The providers at CHCO were initially curious as to whether there is a better way to measure stress on the unit than simply using bed census. Our analysis has revealed that bed census is indeed fundamental to explaining the level of strain on the PICU, which is to be expected considering that the quantity of patients will always have an effect on workload. Our analysis also revealed, though, that admissions are twice as impactful when it comes to explaining strain. Thus, the providers' initial instincts were correct in the sense that other variables significantly impact strain. This insight also suggests that the PICU can most efficiently reduce the level of strain on providers through improvements in the admissions process to reduce or better allocate the level of work involved in admitting a new patient to the PICU.

Considering the effects of the implementation of this stoplight measure is vital in extending its value outside of the PICU. Future research must be conducted to consider the efficacy of predictive staffing based on strain, particularly considering the variables highlighted as primarily indicative of strain (diverts, medication errors, RRT), as described in section 2. Additional analysis must also be conducted to determine the efficacy of the strain metric for predicting events in the future, and adapting this metric to units beyond the CHCO PICU. The stoplight metric and the measures of strain the team identified provide an initial point for predicting levels of strain in the PICU and adjusting staffing accordingly, but must be carefully considered and adjusted as they are implemented in the PICU.

5. References

1. Faddy, M. J., & Mcclean, S. I. (1999). Analysing data on lengths of stay of hospital patients using phase-type distributions. *Applied Stochastic Models in Business and Industry*,15(4), 311-317.
2. Kadri, F., Harrou, F., Chaabane, S., & Tahon, C. (2014). Time Series Modelling and Forecasting of Emergency Department Overcrowding. *Journal of Medical Systems*,38(9).
3. Kellogg, D. L., & Walczak, S. (2007). Nurse Scheduling: From Academia to Implementation or Not? *Interfaces*,37(4), 355-369.

4. Klein, J. D., Beshansky, J. R., & Selker, H. P. (1990). Using the Delay Tool to Attribute Causes for Unnecessary Pediatric Hospital Days. *Medical Care*, 28(10), 982-989.
5. Kolker, A. (2008). Process Modeling of ICU Patient Flow: Effect of Daily Load Leveling of Elective Surgeries on ICU Diversion. *Journal of Medical Systems*, 33(1), 27-40.
6. Rais, A., & Viana, A. (2011). Operations Research in Healthcare: a survey. *International Transactions in Operational Research*, 18(1), 1–31.
7. Teow, KL. (2009). Practical operations research applications for healthcare managers. *Annals of the Academy of Medicine*, 38(6), 564–573.