Proceedings of the 2017 Winter Simulation Conference
W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, eds.

SIMULATING TRIAGE OF PATIENTS INTO AN INTERNAL MEDICINE DEPARTMENT TO VALIDATE THE USE OF AN OPTIMIZATION-BASED WORKLOAD SCORE

Joseph Agor
Kendall McKenzie
Maria E. Mayorga
Osman Ozaltin

Department of Industrial and Systems Engineering
North Carolina State University
Raleigh, NC 27695, USA

Riddhi S. Parikh
Jeanne Huddleston

Department of Medicine
Mayo Clinic, 200 1st Street SW
Rochester, MN 55905, USA

ABSTRACT

This study describes a simulation model that was used to evaluate a proposed workload score. The score was designed to assist in triaging patients into the hospital services of the Division of Hospital Internal Medicine at Mayo Clinic in an effort to more equitably balance workload among the division's provider teams (or services). The first part of this study was the development of a score, using Delphi surveys, conjoint analysis, and optimization methods, that accurately represents provider workload. A simulation model was then built to test the score using historical patient data. Preliminary simulation results reported the proportion of time that each provider team spent working at or above "maximum utilization," as defined by Mayo Clinic experts. The model yielded a 12.1% decrease (on average) in the proportion of time provider teams spent at or above maximum utilization, while simultaneously displaying a more balanced workload across provider teams.

1 INTRODUCTION

At Mayo Clinic in Rochester, MN, patients arrive daily needing general inpatient care provided by the Division of Hospital Internal Medicine (HIM) providers. When a patient arrives, he or she is immediately assigned to a provider team (also called a service), which consists of one doctor, one nurse practitioner or physician's assistant, and one clinical administrative assistant. This provider team is responsible for servicing all the needs of this patient until he or she is discharged from the hospital. However, the decision-making process for determining which provider team will be assigned to which patients is sub-optimal. In fact, it often results in ill-will between provider teams and decreased employee satisfaction. Currently, an incoming patient's assignment is heavily influenced by the current team census. The provider team with the lowest patient census (i.e. the fewest number of patients on their list at that moment in time), but this leaves ample room for imbalanced workload between teams. The team with the fewest number of patients does not necessarily equate to having the lightest existing workload. They may have more efficiently discharged patients that day, increasing the churn of workload, or their patients may be particularly ill and require more attention from the care team.

The goal of this research is to propose a workload score that can be calculated for each provider team in real-time and test it using simulation. The score should accurately represent the amount of work each team is currently experiencing, so that assignment decisions can be made accordingly. By assigning incoming patients to the provider team with the lowest workload score instead of the lowest patient census, we hypothesize that patient workload will be more equitably balanced across HIM provider teams.

The development of this workload score was motivated by a similar score that exists in the Mayo Clinic Emergency Department (ED) and is termed the *overnight score*. To create the overnight score, Mayo Clinic experts identified important factors that were associated with ED workload, and each of the factors was weighted based on provider input; the score is simply the weighted linear combination of these factors. The overnight score is currently calculated in real-time and is used by ED personnel to determine whether or not the daily swing consultant needs to extend his or her shift past 3am. There were multiple improvements due to the implementation of this score (Cabrera et al. 2014). Because the overnight score in the ED has performed well, and because Mayo Clinic experts were receptive to it, we aimed to develop a similar score for the HIM department.

The paper is organized as follows. Section 2 reviews the relevant literature pertaining to the definition, quantification, and balancing of workload in health care settings. Section 3 presents the methods and results that were used to develop the HIM workload score. Section 4 describes the simulation model developed to test the proposed workload score and presents simulation results. Section 5 discusses the implications of the preliminary simulation results.

2 RELEVANT LITERATURE

2.1 Definition and Quantification of Workload in Health Care Settings

Since 1990, the healthcare literature has seen a substantial increase in publications regarding the workload experienced by healthcare professionals. Research has shown that the amount of workload placed on nurses directly affects patient outcomes (e.g. survival or death) (Duffield et al. 2011), as well as nurse satisfaction and resilience in the workplace (Gurses et al. 2009). In order to prevent negative consequences associated with high workload, methods must be created to manage workload. Necessarily, the initial steps in developing a method for managing workload are, first, to define workload and, then, to quantify it. However, a review of the healthcare literature pertaining to workload shows that the medical community has not reached a consensus on the definition and quantification of workload.

There has been a substantial amount of research conducted on nurse workload and staffing. Since nurses are shift-based employees with a limited scope of tasks, this is a natural application area for those researching workload. Nursing workload has been defined in many ways, often by first determining a set of important factors that affect it. For example, Myny et al. (2012) performed a cross-sectional study to identify the factors, outside of patient acuity, that contribute to nurse workload, including high numbers of work interruptions, high patient turnover rate, and high numbers of mandatory government registrations. A plethora of approaches exist for quantification of nurse workload. Kwiecien et al. (2012) reviewed tools used for quantifying nurse workload in the ICU and classified them into five groups, which included patient classification, the Therapeutic Intervention Scoring System-28 (TISS-28), the Nine Equivalents of Nurse Manpower Use Score (NEMS), the Nursing Activities Score (NAS), and experimental methods.

Physicians workload has also been defined and quantified in the literature, usually pertaining to either (i) ED/ICU physicians or (ii) primary care providers / general practitioners. Gedmintas et al. (2010) employed the Australian Triage Scale to develop a tool for managing staffing requirements and understanding resource use in the ED. Levin et al. (2006) also tracked ED physician workload using a human factors approach that included time-motion task analysis and load index. Doerr et al. (2010) used electronic health records and time studies to measure the time and complexity involved in the workload tasks of primary care physicians between their patient visits. While this research can serve as a foundation for defining and quantifying the workload of hospital providers, it is not a direct representation of the daily tasks and decisions experienced in an internal medicine environment.

2.2 Methods for Balancing Workload in Health Care Settings

Over the past few decades, the problem of balancing workload equitably among healthcare providers has been steadily gaining attention. However, few recommendations exists that are tailored to hospital workload. Researchers have primarily focused on developing methods for (i) determining staffing needs based on workload requirements and (ii) distributing workload equitably across providers at key decision points.

Many studies in engineering and management disciplines look at workload for the purpose of staffing. Bard and Purnomo (2005), for example, developed a methodology for nurse scheduling that has the ability to dynamically adjust hospital-wide staffing recommendations based on supply and demand considerations. Additionally, Thorwarth et al. (2009) created a simulation model to represent the dynamic workload experienced by ED providers for use by health care workload management personnel. Punnakitikashem et al. (2013) took a stochastic integer programming approach to solve nurse staffing and assignment problems. The objective of the stochastic program was to minimize both excess nurse workload and staffing costs. Other methods for balancing workload in health care settings focus on key decision points, such as admission decisions and patient allocation decisions. Tseytlin (2009) developed a queueing model for the process of routing patients from the ED to internal wards. The author searched for routing policies that resulted in fairness and good operational performance. Hulshof et al. (2016) used approximate dynamic programming to create a robust framework for allocating new patient admissions to health care resources. Finally, Brewerton (2015) discussed the lack of standardized, robust approaches for achieving equitable workloads across health care teams. The literature is saturated with nurse-focused publications, but is lacking in specifics about physicians, especially as it pertains to hospitalists because this is a relatively new role for physicians.

3 WORKLOAD SCORE DEVELOPMENT

3.1 Delphi Survey and Conjoint Analysis

In order to identify the broad categories that contribute most to HIM provider team workload, a Delphi survey was conducted among the Mayo Clinic providers. Beginning in March of 2015, focus group sessions were held to create an exhaustive list of categories that providers perceived as affecting their workload. Throughout April and May of 2015, more than 900 comments were reviewed and condensed into a set of recurring themes. Nine broad categories were identified: patient churn, lack of autonomy, work interruptions, non-clinical responsibilities, uncertainty about end of day timing, complexity of patients, work inefficiency, changing team members, and geographic location of patients.

To assess the relative impact of category on provider workload, an online survey was created and disseminated to Mayo Clinic providers in December of 2015. The survey asked providers to rank the nine categories in the order in which they contributed to workload (i.e. rank each of the categories from 1 to 9, where 1 identifies the category that contributes most heavily to workload and 9 identifies the category that contributes least to workload). A total of 59 providers responded to the survey, and the distribution of ranks given to each category is shown in Figure 1. The mean rankings of each category (in decreasing order of contribution to workload) were 3.34 for patient churn, 4.27 for work interruptions, 4.37 for work inefficiency, 4.64 for uncertainty about the end of work day, 4.69 for patient complexity, 5.37 for geographical location of patients, 5.76 for changing team members, 6.39 for lack of autonomy, and 6.39 for non-clinical responsibilities.

It was decided that the four categories with the lowest mean rankings (indicating highest average contribution to workload) should be incorporated into the workload score. These were patient churn, work interruptions, work inefficiency, and uncertainty about the end of work day. After further discussion with Mayo Clinic providers and data experts, however, the fourth category (i.e. uncertainty about end of work day) was determined to be too difficult to quantify in a score. Thus, we replaced that category with the patient complexity, which could be easily quantified using diagnosis codes and was next in the ordered list of average rankings.

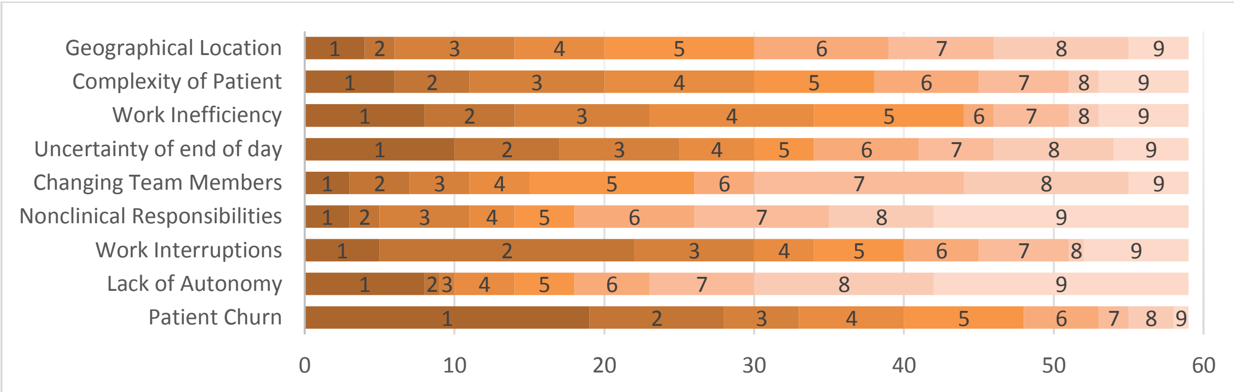

Figure 1: Distribution of rankings of 9 categories contributing to provider workload, as identified using Delphi methods; width of each bar segment is proportional to the number of physicians who assigned ranks 1-9 to each category; 1 identifies the heaviest and 9 identifies the lightest influence on workload.

In order to develop a score to represent workload, the four primary categories were broken down into ten quantifiable factors (see Table 1) that would be reasonable to pull from the Mayo Clinic data systems. All factors have the ability to be updated either daily or in real-time.

Table 1: Ten quantifiable factors used to represent the primary four categories that contribute to provider workload. All factors have the ability to be updated either daily or in real-time. Work inefficiency is represented by factors that indicate future work but require providers to wait for work to arrive.

Patient Churn	Patient Complexity	Work Interruptions	Work Inefficiency
# admissions	# Level 1 patients (low)	# patients with family in town	# admissions assigned but not yet arrived
# discharges	# Level 2 patients (medium)		
# patients (census)	# Level 3 patients (high)	# behavioral patients	# ED registrations

To create a workload score using the ten factors in Table 1, we needed to assign appropriate weights to each factor. To accomplish this, we constructed a secondary survey for the HIM providers, using a choice-based conjoint analysis design. The survey was constructed using XLSTAT software, which used a D-optimal design to optimize the statistical significance produced by the survey results. In the survey, providers were presented with 15 comparison questions formatted as shown in Figure 2. The providers were asked to compare two potential scenarios, each with different levels (i.e. high, medium, low) of the four categories selected for inclusion in the workload score: patient churn, work interruptions, patient complexity, and work inefficiency. There were 19 respondents to our survey, and conjoint analysis was used to elicit relative utilities for each level of the four categories.

Which is worse?	Churn	Patient Complexity	Number of Interruptions	Amount of Indirect Work
Scenario 1:	High	Low	Medium	High
Scenario 2:	High	Medium	Low	High

Figure 2: Example of comparison question presented in choice-based survey. Providers identified the scenario with higher workload in 15 such comparisons, and conjoint analysis was used to elicit utility values for each level (i.e. high, medium, low) of the four broad categories.

Next, we generated 1000 possible combinations of the ten workload factors in our score, using expert opinion to inform the allowable values for each factor. The values of all factors that represented a given broad category were given equal weight and added together, producing a single number to represent each of the four categories. Based on expert opinion, we defined ranges of values that constituted high, medium, and low levels of each category, allowing us to aggregate the 1000 ten-factor combinations into a corresponding representation comprised of only the broad categories. We then assigned a utility to each of the 1000 combinations by weighting its categories by their associated utility values and summing them. These steps resulted in a list of 1000 possible situations (i.e. ten-factor combinations), each with an associated utility value representing the workload that the situation might imply. Situations with higher utility values were considered to have heavier workloads, and those with lower utility values were considered to have lighter workloads. These situations were used as inputs in the optimization model described in Section 3.2.

3.2 Optimization Model

A linear optimization model was created to find the optimal weights for each of the ten factors included in the workload score. Essentially, the optimization model aimed to (i) minimize the number of situations that deviated from the ordered grouping obtained from the conjoint analysis results and (ii) keep the total of the factor weights in each broad category as close as possible to the relative category rankings from the Delphi survey results. The optimization model was trained using 1000 generated observations (i.e. ten-factor combinations) . Each observation consisted of ten factors and was assigned a utility score based on the conjoint analysis results. Observations with identical utilities were grouped, and the groups were ranked in order from heaviest workload (highest utility) to lightest workload (lowest utility).

Let $\{C_u\}_{u=1}^{n}$ be a sequence of sets, e.g. categories that contribute to provider workload, where this sequence is in order of importance (i.e. set C_u has less utility than set C_v for each $u < v$). Let R be the number of groups of situations determined from the conjoint analysis (Section 3.1). Suppose that the groups are ordered from heaviest workload to lightest workload (i.e. the ten-factor combination in group k_1 represents a higher workload situation than that of group k_2, for each $k_1 > k_2$). We define the vectors $\boldsymbol{x}, \boldsymbol{w} \in \mathbb{R}^{\Sigma_u |C_u|}$ as the input observation vector and decision variable vector, respectively. The elements in $\boldsymbol{x}$ and $\boldsymbol{w}$ represent the numerical observations and weights on the elements within the sets $\{C_u\}_{u=1}^{n}$, respectively. We also define the decision variable e_{ij} to represent the error in workload score deviation when comparing observation $i \in G_k$ to observation $j \in G_{k+1}$ for $k = 1, \ldots, R-1$. Given an observation vector $\boldsymbol{x} \in \mathbb{R}^{\Sigma_u |C_u|}$, our goal is to find weights for the $\sum_{u=1}^{n} |C_u|$ factors that result in a workload score calculation that accurately reflects the workload being experienced at the time when this observation is taken. We define our workload score, denoted S, as a weighted linear combination of these factors, $S = \boldsymbol{w}'\boldsymbol{x}$, such that S provides a numerical representation of workload at the time that observation $\boldsymbol{x}$ is taken. The linear program used to determine the weight vector $\boldsymbol{w}$ is as follows:

$$\text{Min} \quad \sum_{k=1}^{R-1} \sum_{i \in G_k} \sum_{j \in G_{k+1}} e_{ij} \tag{1a}$$

$$\text{subject to} \quad \mathbf{w}'\mathbf{x}_i + e_{ij} \geq \mathbf{w}'\mathbf{x}_j \qquad i \in G_k, j \in G_{k+1}, k = 1, \ldots, R-1 \tag{1b}$$

$$\sum_{z \in C_u} w_z \leq \sum_{z \in C_{u+1}} w_z \qquad u = 1, \ldots, n-1 \tag{1c}$$

$$\boldsymbol{w} \geq \mathbf{1}, \ \boldsymbol{e} \geq \mathbf{0} \tag{1d}$$

The objective function (1a) attempts to minimize the total error (i.e. deviation) from the results of both the Delphi survey and the conjoint analysis. The constraints in (1b) ensures that the score preserves the order of the groups generated from the conjoint analysis survey. The constraints in (1c) ensure that the sum of the weights of the elements within each broad category (e.g. patient churn, patient complexity, work interruptions, and work inefficiency) are in order with respect to their average rankings from the Delphi survey. The error term e_{ij} is added to constraint (1b) since preserving the ordered grouping may be infeasible (i.e. there may not exist a weight vector $\boldsymbol{w}$ such that all these constraints are satisfied).

3.3 Resulting Workload Score

We run an instance of the optimization model in the context of our problem, where there are four categories in the sequence of categories, namely, patient churn (C_1), patient complexity (C_3), work interruptions (C_2), and indirect work (C_4). The number of elements in C_1, C_2, C_3, and C_4 are three, three, two and two, respectively, yielding $\boldsymbol{w}, \boldsymbol{x} \in \mathbb{R}^{10}$. The optimal weight vector $\boldsymbol{w}$ found after running the optimization model is shown in Table 2. To test how these weights performed, a second set of 1000 ten-factor combinations was created for use in validation. Using the same procedure as described in Section 3.1, we again assigned utility values to each of these observations and placed them in an ordered grouping. The performance measure used to evaluate the weight vector was the percentage of observations that were misplaced, i.e. did not fall into the ordered grouping correctly once the workload score was calculated. Thus, using the weights in Table 2, we output workload scores for the first 1000 ten-factor combinations in the training set, as well as the next 1000 ten-factor combinations in the validation set. We counted the number of observations that were not in their correct group locations and found that approximately 18.1% and 18.7% of the observations were misplaced in the training and validation sets, respectively. Our conclusion was that this score accurately captures the work being experienced within a medical service and that we could take this score to the next phase of verification through simulation.

Table 2: Optimal weights for each workload score factor, as generated by optimization model. The weighted linear combination of these factors produces a resulting workload score.

Factor Weight	Factor Variable	Factor Description	Category
1.0	w_1	# admissions	Patient Churn
3.5	w_2	# discharges	
2.5	w_3	# patients (census)	
1.0	w_4	# Level 1 patients (low-complexity)	Patient Complexity
2.0	w_5	# Level 2 patients (moderate-complexity)	
3.5	w_6	# Level 3 patients (high-complexity)	
2.5	w_7	# patients with family in town	Work Interruptions
4.0	w_8	# behavioral patients	
1.0	w_9	# admissions assigned but not yet arrived	Indirect Work
5.5	w_{10}	# ED registrations in last hour	

4 SIMULATION MODEL

4.1 Model Structure

A simulation model was built in SIMIO 8 to represent the movement of patients and their associated workloads throughout the Mayo Clinic's HIM department. We modeled 11 provider teams (i.e. Med services), including services 1-4 (teams of multiple medical residents), services 5-9 and 11 (regular provider teams consisting of an MD, nurse practitioner or physician's assistant, and clinical assistant), and service

14 (a single provider who focuses more on doing rounds with existing patients than admitting new patients). The providers on each team were modeled as resources that needed to be seized to complete work.

Patients arrive to the system from six distinct sources, each with different distributions for diagnosis complexity, which was determined by DRG diagnosis codes, and admission processing time. Three entity types were used to represent high, moderate, and low patient complexity, defined by the highest quartile, middle two quartiles, and lowest quartile of possible DRG code weights, respectively. The amount of daily work generated by each patient was different for each complexity level. Other patient attributes assigned upon creation included source location, transfer patient flags, arrival day and time, discharge day and time, etc.

Upon arrival, patients are sent to a decision node that represents the Medical Officer of the Day (MOD), who determines the provider team, or service, to which each patient will be assigned. After assignment to a service, a delay occurs (based on arrival source) to represent the amount of time usually required for the patient to physically arrive at the Mayo Clinic HIM department. After this delay, the patient entity is considered to be under the care of its assigned provider team until it is either transferred to a different provider team or discharged entirely from the hospital. Each service is modeled as a sub-model in the simulation, where the providers are modeled as workers that can be seized by jobs created by patients each day. A daily check within the simulation determines when each patient is scheduled to exit the sub-model (according to historical data). Work, in the form of multiple different job types, is produced on each day that a patient remains in the sub-model. Jobs (e.g. rounds, family visits, paperwork, etc.) are created at the start of each day, and the provider resources within the sub-model work to complete these jobs. A large portion of the work created deals with patient admissions and discharges. Figure 3 shows a high-level visualization of the simulation model. A sample sub-model flow diagram is displayed in Figure 4.

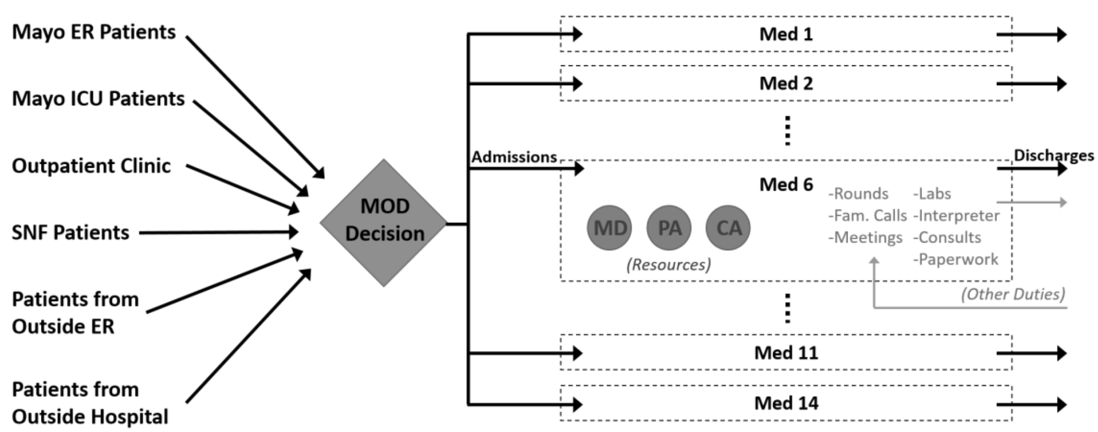

Figure 3: High-level flow diagram of simulation model. Patient types arrive from six sources, each with different distributions for diagnosis complexity and admission processing time. Each patient is assigned to a HIM provider team (i.e. Med service) by the Medical Officer of the Day (MOD). Each Med service is a sub-model that seizes resources (i.e. MD, physician's assistant, clinical assistant) who act as servers for generated tasks, such as admission work, rounds, paperwork, etc. Utilization of these resources is tracked.

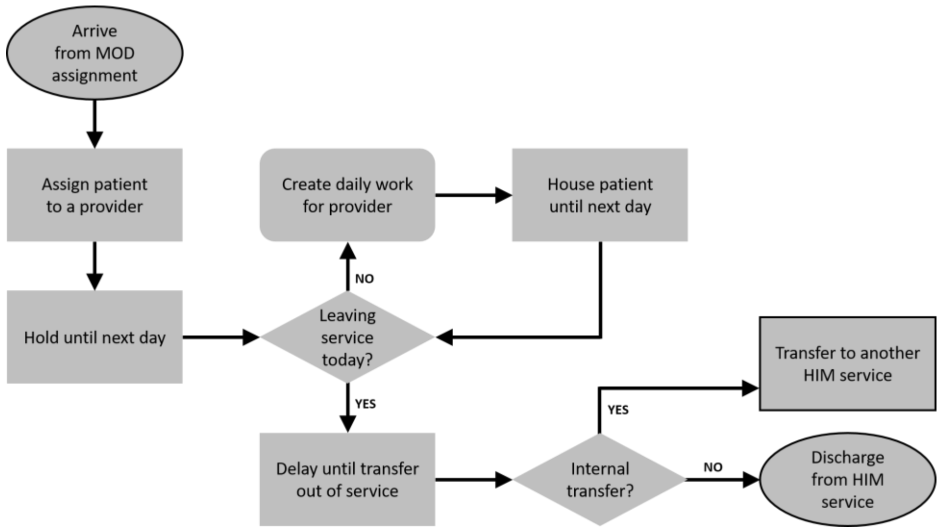

Figure 4: Flow diagram of a single provider team (i.e. Med service) sub-model. Outlined modules identify the entrance and exit points for the sub-model. All patients are assigned to a provider team by the MOD, thus entering the sub-model. Patients exit the sub-model by either (i) being transferred to another Med service sub-model (i.e. changing provider teams) or (ii) being discharged entirely from the HIM department.

The simulation model samples from historical data in some cases, and it makes select assumptions in others. For example, patient arrival and discharge locations, arrival and discharge times, and patient complexity were matched to historical data. Distributions for provider job processing times, the number of behavioral patients, and the number of patients with families in town were developed from expert opinion. Using the historical data, we were able to first simulate the historical assignment locations of patients and track the resulting resource utilization. We were then able to re-run the model using our workload score to make patient assignment decisions and compare the resulting resource utilization with that of historical assignment policy.

4.2 Simulation Results

The simulation was run using data from January 1, 2013 through December 31, 2015. Since we did not have data prior to this time period, the model was run with a one-year warm-up period to allow the total patient census across all HIM services to stabilize. After a one-year warm-up period, we began collecting data for calculating results.

The experts at the Mayo Clinic evaluating the workload score were interested in comparing simulation output relating to a specific metric: the proportion of days per month that each service spent any amount of time working at *maximum utilization*. The term *maximum utilization* was defined differently across provider teams. Provider teams in HIM services 1-4 were considered to have reached maximum utilization on any day that either (i) their census hit 12 or (ii) all of their admission slots were used. Provider teams in HIM services 5-9 and 11 were considered to have reached maximum utilization on any day that their census hit 14. The provider team in HIM service 14 was considered to have reached maximum utilization on any day that its census hit 12.

Figure 5 shows the proportion of time per month that each provider team reached its maximum utilization (on average) in the simulation. Using the proposed workload score to make patient assignment decisions resulted in a 12.1% decrease (on average) in the proportion of time per month provider teams spent at or above maximum utilization. The proportions in Figure 5 are calculated by averaging the monthly maximum utilization proportions across the entire 24-month observation period.

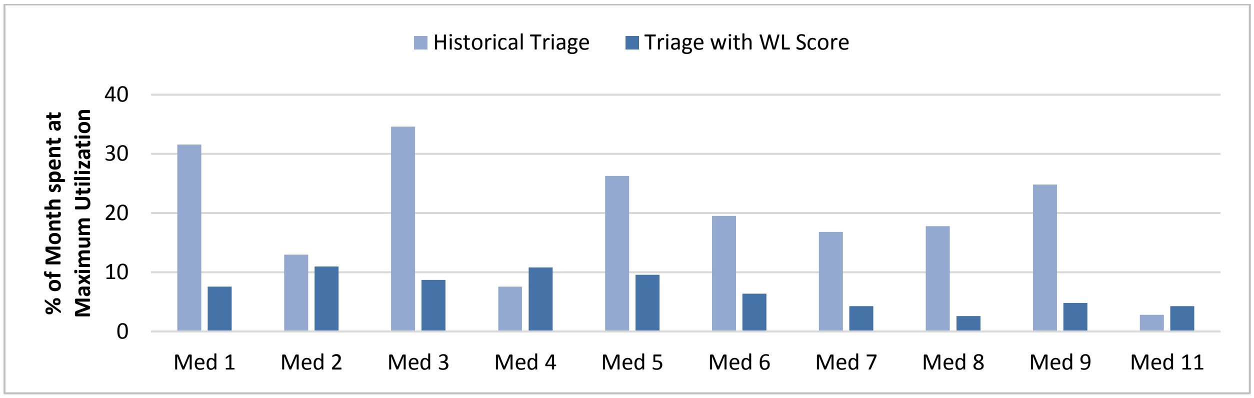

Figure 5: Proportion of month each Med service reached maximum utilization, as defined by Mayo Clinic experts. Provider teams in services 1-4 were considered to have reached maximum utilization on any day that either (i) their census hit 12 or (ii) all of their admissions slots were used. Provider teams in services 5-9 and 11 were considered to have reached maximum utilization on any day that their census hit 14.

The HIM leadership was interested in understanding the difference in utilization between non-resident provider teams (HIM services 5-9, 11, and 14) and resident provider teams (HIM services 1-4). Figures 6 and 7 provide comparisons of the simulated results for these two groups under historical triage assignments and triage assignments using our proposed workload score. Figure 6 shows the proportions of days in the month that two non-resident services (i.e. Med 5 and Med 7) reached a census level greater than 14. There is improvement when patients are triaged using the proposed workload score with respect to time spent at maximum utilization, along with the more desired outcome of more equitably-balanced workload between the two services throughout the displayed time window. Figure 7 shows distribution of the proportions of time the resident services, which must follow strict rules that prevent them from taking on too much work, spent *below* maximum utilization throughout the months of data collection. It is clear that the proportion of time that the resident services are *not* fully utilized decreases when patients are triaged using the proposed workload score. This is an improvement that HIM providers are looking to achieve, since there is no risk of overworking the resident teams. Again, Figure 7 also shows improvement in the balance of workload across the resident services when the proposed workload score is used to make assignment decisions instead of the current census.

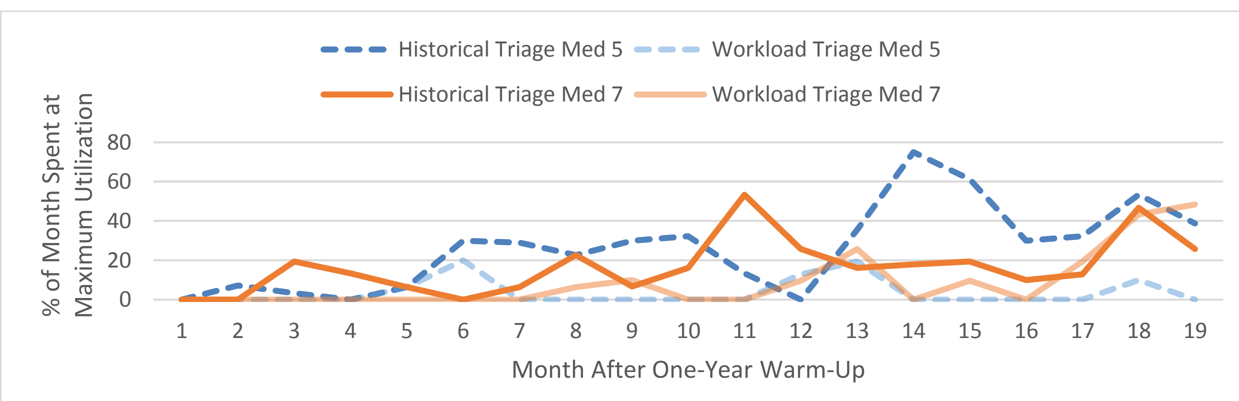

Figure 6: Proportion of days per month that census level was at least 14 patients for Med services 5 (dashed line) and 7 (solid line) when comparing historical triage (darker color) and triage by simulated workload score (lighter color).

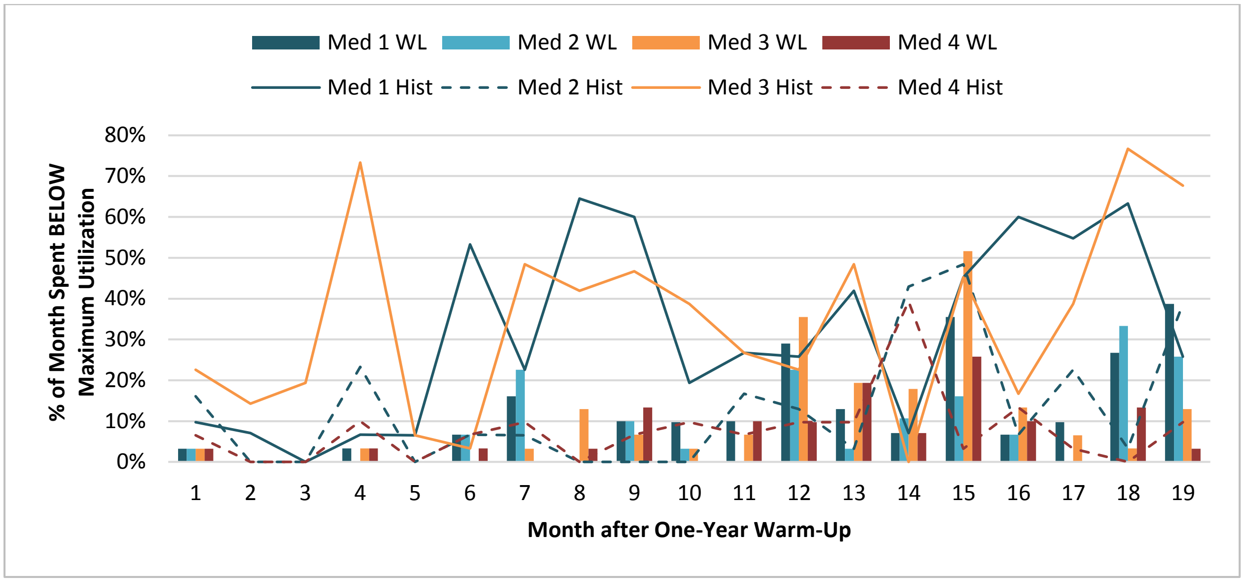

Figure 7: Proportion of days per month that resident services 1-4 did *not* reach maximum utilization. Lines represent historical triage, and bars represent simulated triage by our workload score.

5 DISCUSSION AND CONCLUSION

Through the use of a combination of methodologies (i.e. Delphi survey, conjoint analysis, optimization), we were able to create a workload score that proved successful in our simulation model. The measure that Mayo Clinic staff found most informative for workload comparison was an interesting one: the percentage of days each month that a care team reached maximum utilization. Our simulation yielded a 12.1% decrease (averaged across all teams) in this metric when the proposed workload score was used to determine the provider team that should receive each incoming patient, instead of simply assigning incoming patients to the team with the fewest patients currently in their care. The results also show a significant decrease in the variation between team workloads when our workload score is used for patient assignment decisions. Furthermore, the provider teams staffed by medical residents showed a reduction in time spent being under-utilized, which is an improvement that Mayo Clinic HIM management was hoping to achieve. This confirms that the proposed workload score has the potential to balance workload more equitably across provider teams.

Not only did we achieve a more equitable workload between HIM care teams, but we were also able to provide a workload score calculation that more accurately represents employees' perceptions of their own workloads. The score can be implemented by pulling only ten numbers from the hospital data systems. In the future, we hope to implement this score within the Mayo Clinic and test its performance. Our next steps are to refine the score with provider input and implement this workload score as a triaging tool within the Mayo Clinic system for a beta test.

The methodology used in the development of this workload score for physicians could be extended to develop scores to measure workload amongst nurses where the score can influence nurse staffing needs in real time. Outside of healthcare, this methodology could be used to develop workload scores for the public school system, to assist \administrators in the allocation of resources (i.e. funding, technology, teachers, etc.) amongst schools within their district or state based on the workload being experienced. The proposed methodology provides a systematic way to develop a score that represents a significant amount of information pertaining to workload so that decision makers can make well informed decisions without having to look through great amounts of data.

ACKNOWLEDGMENTS

This work was supported by the Sepsis Early Prediction Implementation System (SEPSIS) Collaborative, funded by NSF, SCH grant no. 1522107; a Mayo Clinic Department of Medicine Inpatient Innovation Award; and the Division of Hospital Internal Medicine. Many thanks to Mr. Lynn Loynes, who facilitated most of the focus groups leading to the categorical themes used in this simulation model.

REFERENCES

Bard, Jonathan F, and Hadi W Purnomo. 2005. "Hospital-Wide Reactive Scheduling of Nurses with Preference Considerations." *IIE Transactions* 37 (7): 589–608. doi:10.1080/07408170590948468.

Brewerton, Amy. 2015. "Achieving Equitable Workloads." *Community Practitioner* 88 (6): 13.

Cabrera, Daniel, Jeffrey L. Wiswell, Vernon D. Smith, Andy Boggust, and Annie T. Sadosty. 2014. "A Novel Automatic Staffing Allocation Tool Based on Workload and Cognitive Load Intensity." *American Journal of Emergency Medicine* 32 (5). Elsevier Inc.: 467–68. doi:10.1016/j.ajem.2013.12.051.

Doerr, Emmett, Kevin Galpin, Cedrella Jones-Taylor, Steven Anander, Charles Demosthenes, Susan Platt, and Sanjay Ponkshe. 2010. "Between-Visit Workload in Primary Care." *Journal of General Internal Medicine* 25 (12): 1289–92. doi:10.1007/s11606-010-1470-2.

Duffield, Christine, Donna Diers, Linda O'Brien-Pallas, Chris Aisbett, Michael Roche, Madeleine King, and Kate Aisbett. 2011. "Nursing Staffing, Nursing Workload, the Work Environment and Patient Outcomes." *Applied Nursing Research* 24 (4). Elsevier Inc.: 244–55. doi:10.1016/j.apnr.2009.12.004.

Gedmintas, Audra, Nerolie Bost, Gerben Keijzers, David Green, and James Lind. 2010. "Emergency Care Workload Units: A Novel Tool to Compare Emergency Department Activity." *Emergency Medicine Australasia* 22 (5): 442–48. doi:10.1111/j.1742-6723.2010.01322.x.

Gurses, Ayse P., Pascale Carayon, and Melanie Wall. 2009. "Impact of Performance Obstacles on Intensive Care Nurses' Workload, Perceived Quality and Safety of Care, and Quality of Working Life." *Health Services Research* 44 (2P1): 422–43. doi:10.1111/j.1475-6773.2008.00934.x.

Hulshof, Peter J H, Martijn R K Mes, Richard J. Boucherie, and Erwin W. Hans. 2016. "Patient Admission Planning Using Approximate Dynamic Programming." *Flexible Services and Manufacturing Journal* 28. Springer US: 30–61. doi:10.1007/s10696-015-9219-1.

Kwiecień, Katarzyna, Maria Wujtewicz, and Wioletta Mędrzycka-Dąbrowska. 2012. "Selected Methods of Measuring Workload among Intensive Care Nursing Staff." *International Journal of Occupational Medicine and Environmental Health* 25 (3): 209–17. doi:10.2478/S13382-012-0035-5.

Levin, Scott, Dominik Aronsky, Robin Hemphill, Jin Han, Jason Slagle, and Daniel J. France. 2007. "Shifting Toward Balance: Measuring the Distribution of Workload Among Emergency Physician Teams." *Annals of Emergency Medicine* 50 (4): 419–23. doi:10.1016/j.annemergmed.2007.04.007.

Myny, Dries, Dirk Van Goubergen, Micheline Gobert, Katrien Vanderwee, Ann Van Hecke, and Tom Defloor. 2011. "Non-Direct Patient Care Factors Influencing Nursing Workload: A Review of the Literature." *Journal of Advanced Nursing* 67 (10): 2109–29. doi:10.1111/j.1365-2648.2011.05689.x.

Punnakitikashem, Prattana, Jay M. Rosenberber, and Deborah F. Buckley-Behan. 2013. "A Stochastic Programming Approach for Integrated Nurse Staffing and Assignment." *IIE Transactions* 45 (10): 1059–76. doi:10.1080/0740817X.2012.763002.

Thorwarth, Michael, Amr Arisha, and Paul Harper. 2009. "Simulation Model to Investigate Flexible Workload Management for Healthcare and Servicescape Environment." In *Proceedings of the 2009 Winter Simulation Conference*, edited by M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, and R. G. Ingalls, 1946–56. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc. doi:10.1109/WSC.2009.5429210.

Tseytlin, Yulia. 2009. "Queueing Systems with Heterogeneous Servers: On Fair Routing of Patients in Emergency Departments_Application." Israel Institute of Technology. https://iew3.technion.ac.il/serveng/References/thesis-yulia.pdf.

AUTHOR BIOGRAPHIES

JOSEPH AGOR is a PhD student in Operations Research at North Carolina State University and currently a graduate research assistant in the university's Department of Industrial and Systems Engineering. He holds a B.S. in Civil Engineering and a Master's degree in Mathematics, both from the University of Nevada, Reno. His research interests include multilevel optimization, stochastic programming, and game theory, with applications in health care. His e-mail address is jkagor@ncsu.edu.

KENDALL MCKENZIE is a PhD student in the College of Design at North Carolina State University (NCSU), and a graduate research assistant in the university's Department of Industrial and Systems Engineering. She holds a B.A. in Visual Communication from the University of North Carolina, Chapel Hill, as well as a Master's degree in Industrial and Systems Engineering from NCSU. Her research interests combine methods from both engineering and design thinking to solve problems in health care systems. Her e-mail address is kemckenz@ncsu.edu.

MARIA E. MAYORGA is an Associate Professor of Personalized Medicine in the Edward P. Fitts Department of Industrial and Systems Engineering at North Carolina State University. She received her B.S. degree in Mechanical Engineering from The George Washington University, and her M.S. and PhD in Industrial Engineering and Operations Research (IEOR) from the University of California, Berkeley. Her research interests include predictive models in health care and health care operations management. Her e-mail address is memayorg@ncsu.edu.

OSMAN OZALTIN is an Assistant Professor of Personalized Medicine in the Edward P. Fitts Department of Industrial and Systems Engineering at North Carolina State University. He received his B.S. degree in Industrial Engineering from Bogazici University in Istanbul, Turkey, and his M.S. and PhD in Industrial Engineering from the University of Pittsburgh. His research interests span theoretical, computational, and applied aspects of mathematical programming, focusing on multilevel stochastic optimization problems arising in public health policy-making, personalized medical decision-making, and health care delivery. His e-mail address is oyozalti@ncsu.edu.

RIDDHI PARIKH is an Internal Medicine Specialist at the Mayo Clinic in Rochester, MN. She graduated from medical school in 2005 and has 12 years of diverse experiences, especially in Internal Medicine. Dr. Parikh affiliates with Mayo Clinic Hospital Rochester and cooperates with other doctors and specialists in medical group Mayo Clinic. Her e-mail address is parikh.riddhi@mayo.edu.

JEANNE HUDDLESTON is an Associate Professor of Medicine at the Mayo Clinic in Rochester, MN, currently working within the Hospital Internal Medicine Department. She obtained her B.S. and M.D. from Michigan State University, as well as her M.S. in Industrial Engineering form Arizona State University. Her research goal is the innovative translation of systems engineering principles to health care delivery in an effort to improve the safety, efficiency, and reliability of the health care experience for patients, their families, and all providers. She has led the integration of engineering principles and research into clinical care, culminating in the creation of the first Clinical Engineering Learning Laboratory with the Department of Emergency Medicine. Her e-mail address is huddleston.jeanne@mayo.edu.