August 2018

Labor Planning Outcomes: Systemic Management Models, Human Interactions, and Knowledge Sharing

Richard J. Tarpey
University of South Florida, cssqtarpey@gmail.com

Labor Planning Outcomes: Systemic Management Models, Human Interactions,

and Knowledge Sharing

by

Richard J. Tarpey

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Business Administration
Department of Business
College of Business
University of South Florida

Major Professor: Timothy B. Heath, Ph.D.
JoAnn F. Quinn, Ph.D.
James Stikeleather, D.B.A.
Eric Eisenberg, Ph.D.

Date of Approval:
September 18, 2018

Keywords: Staffing, Scheduling, Forecasting Accuracy, Hospital

Table of Contents

List of Figures ... iii

List of Tables ... iv

Abstract ... vi

Chapter One: Introduction ... 1
- References ... 2

Chapter Two: Precedent Research ... 3
- References ... 4

Chapter Three: Research Method & Structure ... 5
- References ... 7

Chapter Four: Labor Forecasting, Staffing, and Scheduling Outcome Measurement ... 8
- Introduction ... 8
- Motivation ... 9
- Research Design ... 12
- LPAC Phase Performance Measurement ... 13
- Workload Demand Forecasting Metric Selection ... 15
- Schedule Quality Metric Selection ... 16
- Allocation (Staffing) ... 17
- Metric Analysis – Artifact Iteration #1 ... 18
- Metric Analysis – Artifact Iteration #2 ... 22
- Metric Analysis – Artifact Iteration #3 ... 25
- System Based Measurement ... 30
- Conclusions, Limitations, and Future Directions ... 31
- References ... 36

Chapter Five: Case Study-Human interaction management impact on hospital labor planning ... 40
- Note to Reader ... 40
- Introduction ... 40
- Human Interaction Management ... 43
- The Case ... 45
- Measurement Criteria ... 45
- Current State Background ... 47
- New Model Design ... 52
- Forecasting and Scheduling Process Changes ... 52
- Discussion ... 54
- Conclusion and Next Steps ... 58
- References ... 59

Chapter Six: Hospital Labor Planning and Allocation: A System Out of Balance ... 61
- Introduction ... 61
- Conceptual Framework ... 63
- Method ... 66

The Hospital as a System 67
Labor Forecasting, Scheduling, and Staffing Subsystem 69
LPAC Management Model Development 71
Results Summary...75
Workload Forecasting 77
Scheduling...79
Staffing..84
Discussion 86
Limitations and Next Steps 88
References 89

Chapter Seven: Overall Conclusion..92

Chapter Eight: Opportunities for Future Research 95

Appendix 1 - Accepted Metric Calculation Formulas 97

Appendix 2 – Muma Business Review Reproduction Permission...99

List of Figures

Figure 1: Study Objectives 7
Figure 2: Integrated Forecasting, Scheduling, and Staffing (LPAC Cycle) 10
Figure 3: The Complexity of Labor Planning and Allocation Within a Hospital 11
Figure 4: Forecast Patients, FTE, and Bin Determination 23
Figure 5: Actual Patients, FTE, and Bin Determination 23
Figure 6: Workload Forecasting Control Charts Iteration #3 Results 28
Figure 7: Schedule Quality Completeness Graphs 29
Figure 8: Staffing Control Charts Iteration #3 Results 30
Figure 9: Service Line Forecasting Control Charts 33
Figure 10: Service Line Scheduling Completeness Charts 34
Figure 11: Service Line Staffing Control Charts 35
Figure 12: RAD (Decentralized Model) 49
Figure 13: RAD (Centralized Model) 51
Figure 14: RAD (Centralized Model) –Review Touchpoints 55
Figure 15: Simplified Labor Allocation Stock and Flow Diagram 64
Figure 16: Hospital Organizational Structure System of System Complexity 68
Figure 17: Operating Departments or Service Line Isolation from Higher Layer Systems 69
Figure 18: Labor Planning and Allocation Cycle (LPAC) 70
Figure 19: LPAC Quality & Effort Relationships 71
Figure 20: Example Role Activity Diagram – Forecasting and Scheduling 74
Figure 21: Example Interaction Management Card – Process P1 75
Figure 22: Simplified Causal Loop Diagram (2 Subsystems) 76

List of Tables

Table 1: Initial Workload Forecasting Accuracy Measures 15

Table 2: Initial Schedule Quality Measures 17

Table 3: Initial Staffing Accuracy Measures 18

Table 4: Workload Forecasting Metrics Iteration #1 Results 19

Table 5: Schedule Quality Metrics Iteration #1 Results 19

Table 6: Staffing Metrics Iteration #1 Results 19

Table 7: Iteration #1 Metric Acceptance Results 21

Table 8: Workload Forecasting Accuracy Measures (Iteration #2) 22

Table 9: Workload Forecasting Metrics Iteration #2 Results 24

Table 10: Schedule Quality Metrics Iteration #2 Results 25

Table 11: Staffing Accuracy Measures (Iteration #2) 25

Table 12: Staffing Metrics Iteration #2 Results 26

Table 13: Iteration #2 Metric Acceptance Results 26

Table 14: Workload Forecasting Metric Iteration #3 Results 27

Table 15: Schedule Quality Metric Iteration #3 Results 27

Table 16: Staffing Metric Iteration #3 Results 29

Table 17: Allocation Iteration #3 Results (Service Line) 31

Table 18: Final Metric Acceptance Results 31

Table 19: Schedule Quality Components 46

Table 20: Critical Care Department Staffing Grid Example 47

Table 21: REACT analysis of Actor in Forecaster Role (Pre-Implementation) 50

Table 22: REACT analysis of Actors in Scheduler Role (Pre-Implementation) 50

Table 23: SbMAE Scores for Forecasting Before Implementation 56

Table 24: Schedule Quality Scores Before Implementation 57

Table 25: LPAC Performance Metric Acceptance Criteria 66

Table 26: Example Staffing Grid 78

Table 27: Workload Forecasting Results 80

Table 28: Scheduling Results 83

Table 29: Staffing Results 85

Abstract

This project takes a systemic approach to hospital labor planning and allocation rather than sub-optimizing the individual components of workload demand forecasting, scheduling, and staffing separately. The research considers all three components within their interdependent, dynamic, cyclical systemic nature to develop a better labor planning and allocation cycle (LPAC) management model across the various subsystems of the hospital. We used an Action Design Research (ADR) method to the guided emergence of innovative artifacts – Systemic LPAC Management Model and LPAC Performance Metrics – that we evaluate and improve through interventions in situ with practitioners. The Systemic LPAC Management Model leveraged an optimization of organizational structures, work tasks and human interactions based on patient flow to create improved outcomes. Outcomes were measured via the LPAC Performance Metric artifact to assess pre and post-implementation performance. The ADR research method allowed us to assess the resulting utility and acceptance of the new model and metrics in a real-world hospital environment from both a qualitative and quantitative perspective. Implementation of the new model resulted in outcome improvement in each of the individual LPAC phases. Additionally, we observed labor management flexibility and responsiveness improvement due to the systemic approach of improving upon the previous siloed and narrowly focused labor-management model.

Chapter One: Introduction

The labor planning and allocation cycle (LPAC) in hospitals is more complex and less predictable than other industries requiring more collaborative and adaptive solutions. Hospital organizational structures cause much of this complexity due to the intricate systems formed from a collection of inter-related departments interacting within several sub-systems (service lines) in several aspects of operations (patient flow, materials flow, pharmaceutical flow, labor, etc.). While a system perspective is used to view much of operations, rarely is this view applied to labor planning and allocation. Hospitals typically focus on departmental labor performance results leading departmental leaders to be narrowly focused and siloed causing them to overlook system efficiency opportunities in favor of sub-optimization. Tribal-like cultures can take hold forming departmental protectionist environments disconnecting the system approach. This departmental focus, however, can hinder the understanding of the many interconnections within the system that are critical success factors in labor planning. A system understanding is a key to ensuring critical information flows across resources as well as identifying key intervention points where action can be applied to guide successful outcomes.

To address this disconnected system, our research moved beyond an inter-departmental and centralized policy approach to the conduct of the Labor Planning and Allocation Cycle (LPAC) within a hospital to take a systemic approach. The project explored the sub-optimization of an intra-departmental focus that seeks to optimize the individual components of workload forecasting, scheduling, and allocation separately to compare against a systemic approach considering all three interconnected components. This complex nature of workload optimization – the right nurse at the right time for the right level of patient acuity – is impacted by the LPAC constructed from a collection of departments interacting within a larger system.

We designed a new LPAC management model artifact leveraging a systemic optimization of organizational structures, component interactions, knowledge sharing, and human interactions based on patient flow to create improved outcomes. We used an Action Design Research (ADR) (Mullarkey & Hevner, 2015) method in three separate phases to guide the emergence of innovative artifacts – the Systemic LPAC Management Model and LPAC Performance Metrics – that we evaluated and improved through interventions in situ with practitioners and implemented within real-world hospital environments. The project compared outcomes under the new LPAC management model with pre-implementation outcomes highlighting the advantageous improvements realized through systems management as applied to labor planning and allocation.

References

Mullarkey, M. T., & Hevner, A. P. (2015). Entering Action Design Research. In Donnellan B (Ed.), *DESRIST* (pp. 121–134). LNCS 9073.

Chapter Two: Precedent Research

There are three main gaps in academic research we attempted to address. One is that research has typically investigated the forecasting, scheduling, and staffing problems as separate issues attempting to propose narrowly focused solutions tested in limited problem set environments. This approach has not been very helpful in providing an understanding of labor planning and allocation across an organization. Having a better methodology for just one component such as labor scheduling provides no guarantee that outcomes (employee satisfaction or financial) will improve after the staffing function since poor staffing function performance can override good scheduling function performance. Outcomes are dependent on the interaction of all of the system and sub-system components of labor planning and allocation from workload forecasting through labor scheduling to labor staffing. Work within each function has the possibility of negatively or positively affecting intermediate and end-stage outcomes by correcting or overriding performance in a previous function. Prior researchers have not significantly explored the interconnectivity of these concepts.

The second gap exists where academic research has limited its focus to single environments such as emergency rooms, critical care, or medical inpatient departments for studying problems. The "system" nature of hospitals causes department patient flows to interact and potentially affect other departments' workloads. Staff also can flow within the system subject to hard and soft constraints as labor is moved between departments during the different LPAC phases to accommodate needs in other departments. Despite the prevalent view that hospitals are complex systems, they still tend to be siloed organizations along departments or service lines where labor is concerned and typically studied from this same fragmented perspective. Academic studies have paralleled this approach by focusing on individual components of the LPAC within isolated areas of a hospital. Rarely have academics studied the system organization as it applies to labor as they have applied this perspective to patient flow or supply chain analysis. A thorough study of labor planning and allocation needs to take into account nursing units in multiple service lines of a hospital to study involved component interactions.

The LPAC tasks of forecasting, scheduling and staffing can be considered knowledge-intensive and therefore are heavily impacted by human interactions, judgments, and communications which are less represented by machine like sequences of activities (Seo, Choi, Kim, & Lee, 2012; Seo, Yoon, Lee, & Kim, 2011). This understanding represents the third gap in academic research. The majority of precedent research approaches each function from an automation goal perspective. Main research objectives have included proposed automated solutions that attempt to remove humans from the process rather than recognizing the role and contribution of human interactions and judgment. These studies have offered narrow solutions focused on individual department functions with very little consideration for the "system" nature of a typical hospital or the complexity and variability of inputs overlooking the value added through human input and judgment. An understanding of the critical knowledge and information to be shared across the necessary resources should be a prerequisite to applying a technology solution, or an organization runs the risk of automating existing bad processes and focusing on the wrong or missing critical information during decision-making functions.

References

Seo, W., Choi, S., Kim, K., & Lee, J. Y. (2012). A procedural framework for dynamic changes of human interactions in knowledge-intensive services. *Expert Systems with Applications*, *39*(3), 2720–2732. https://doi.org/10.1016/j.eswa.2011.08.129

Seo, W., Yoon, J., Lee, J., & Kim, K. (2011). A state-driven modeling approach to human interactions for knowledge-intensive services. *Expert Systems with Applications*, *38*(3), 1917–1930. https://doi.org/10.1016/j.eswa.2010.07.124

Chapter Three: Research Method & Structure

This research sought to determine if a systemic centralized approach to the LPAC functions could provide better outcomes. The following research questions were developed to guide the project:

- *Question #1*: What will a better future look like with a Systemic LPAC deployed in a typical hospital?
- *Question #2*: How will a systems approach driven by resource flows and not limited by departmental structures lead to greater optimization of the balance between nursing labor, nursing satisfaction, and patient outcomes?

The action research components of the project sought to analyze behavior at the individual and group levels. Analysis at the individual level included work and task behavior from each role perspective in support of the overall objective outcomes of the LPAC process. Analysis at the group level included the complex interactions of the system components and the subsequent impact on performance outcomes.

This research proposes that a gap exists between the desire to optimize the LPAC and the actual results of existing LPAC instantiations in hospitals. Technology systems that automate existing tasks fail to account for human judgment, communication, interaction, and sense-making. This project's objective is to document an understanding of the critical inputs to LPAC components; knowledge generated and shared within and between each component, work and task structures that impact knowledge creation and sharing, and organizational structures that facilitate confirming and corrective interventions. We postulate that this understanding can subsequently be leveraged to re-shape the Systemic LPAC model as the framework for designing new work and organizational structures/systems to facilitate outcome improvement. The complexity of this objective required a research method that goes beyond observation and explanation. Consequently, we chose a research method grounded in action research that involves intervention with practitioners in situ to co-create and co-evaluate multiple iterations of a future state Systemic LPAC model that optimized the balance between nursing labor, cost, and patient outcomes.

The project used the four-stage elaborated Action Design Science Research (eADR) model (Mullarkey & Hevner, 2015) to structure the research method. Embedding researchers within the practitioner environment provided the opportunity for iteratively designing and refining artifacts used within the LPAC cycle. Through the iterative use – define, build, evaluate, learn, reflect, re-define process – we found these artifacts became increasingly more relevant and contributed more toward the overall solution.

This research consisted of three phases each documented within a separate paper (refer to Figure 1). The project began by establishing foundational performance metrics for measuring individual outcomes in the areas of demand forecasting accuracy, labor scheduling and labor staffing (Outcome Metrics in Figure 1). Our goal was to assess LPAC phase performance pre and post implementation of a new LPAC management model. To accomplish this goal there needed to be a way of measuring success within each component of the LPAC. No standard set of industry or academic measures was found to suffice that assessed from a departmental and systems perspective. Therefore, these artifacts were adapted from existing documented metrics, evaluated, and improved through interventions in situ with practitioners as detailed in the first paper (Chapter 1) - Labor Forecasting, Staffing, and Scheduling Outcome Measurement. We felt it was critical to use a method involving deep practitioner engagement to sufficiently test the applicability and utility of the metrics in a production environment.

The second phase of the project again utilized action design research to develop a new model of systemic, centralized labor planning and allocation serving to facilitate knowledge sharing, human judgment, and human interactions to leverage system opportunities and improve outcomes (Model Development in Figure 1). This phase documented existing siloed approaches focusing on social (human) interactions and the impact of centralizing and restructuring work on LPAC outcomes. The project thoroughly documented each component of the LPAC concerning roles, communications, social interactions, technical interactions, inputs, and outputs. The objective was the restructuring of roles and activities within the labor forecasting, scheduling and allocation work streams to facilitate a systems approach increasing participant interaction, communication and ultimately knowledge sharing. The project hypothesized that critical knowledge is unlocked via the systems view across silos through work and role restructuring resulting in more optimal communication and knowledge sharing to improve outcomes. The project hypothesized that documentation of structures via role activity and state diagrams provide role participants with guidance in their day to day activities to guide and sustain the model. This ADR project resulted in the development of the Systemic LPAC Management Model detailed in the second paper (Chapter 2) - Case Study: Human interaction management impact on hospital labor planning.

The third phase of the project extended the new model once again via action design research across a larger sample set utilizing a centralized approach to labor planning and allocation across multiple service lines and hospitals to determine the potential impact to outcomes (ADR – Model Extension in Figure 1). This project further evaluated and improved the model through interventions in situ with practitioners and is detailed in the third paper (Chapter 3) - Hospital Labor Planning and Allocation: A System Out of Balance.

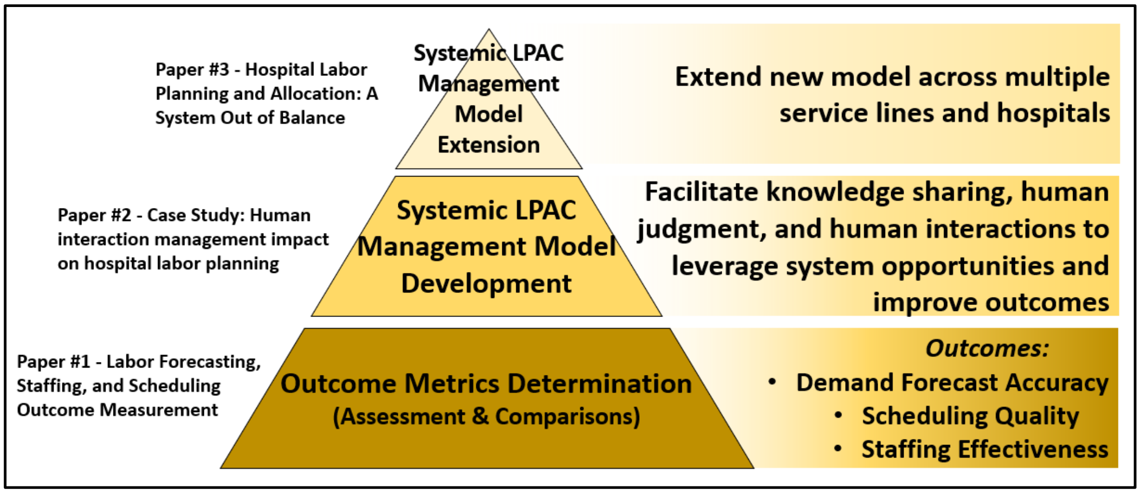


Figure 1 - Study Objectives

References

Mullarkey, M. T., & Hevner, A. P. (2015). Entering Action Design Research. In Donnellan B (Ed.), *DESRIST* (pp. 121–134). LNCS 9073.

Chapter Four: Labor Forecasting, Staffing, and Scheduling Outcome Measurement

Abstract: This project takes a systemic approach to quantifying labor forecasting, scheduling, and staffing outcomes. Moreover, rather than optimizing the individual components of workload forecasting, scheduling, and staffing separately, this research considers all three components within their interdependent, dynamic, cyclical systemic nature. The research objective is to define outcome metrics conducive to performance assessment of each phase of the labor planning and allocation cycle (LPAC) leveraging a systemic optimization of organizational structures, work tasks and human interactions based on patient flow to create improved outcomes. We use an Action Design Research (ADR) method to the guided emergence of an innovative artifact – LPAC Performance Metrics – that we evaluate and improve through interventions in situ with practitioners. The study concludes with a discussion of the resulting utility and acceptance of the development metrics in a real-world hospital environment.

Keywords: Action design research (ADR), Labor management, Labor Scheduling, Labor Staffing

Introduction

Balancing healthcare labor quantity, labor costs, and desired service-quality levels is critical for service success. Research shows that successful labor planning and allocation helps control labor costs, improve employee satisfaction/retention (Harms, 2009; Kuhar, Miller, Spear, Ulreich, & Mion, 2004; Silvestro & Silvestro, 2000; Sjögren, Fochsen, Josephson, & Lagerström, 2005; Yildirim & Aycan, 2008) and facilitate positive patient outcomes (Aiken, 2002; Cimiotti, Aiken, Sloane, & Wu, 2012; Kane, Shamliyan, Mueller, Duval, & Wilt, 2007; Mark, Harless, McCue, & Xu, 2004; Needleman et al., 2011; Needleman, Buerhaus, Mattke, Stewart, & Zelevinsky, 2002). Hospitals typically take an intra-departmental approach to labor planning and allocation. Even if they can optimize the balance for a given department, this approach ignores the larger system within which patients and staff naturally ebb and flow. The resulting lack of communication and knowledge sharing across departments results in narrowly focused optimization, and improvement efforts as department leaders limit their focus and miss the bigger picture associated with having the right staff at the right place at the right time given a real-time patient, not artificially established organizational demands. Previous studies have demonstrated the possibility that a more inter-departmental focus may lead to improvement opportunities. Two specific studies found that a hospital-wide nurse staffing process, nurse organizational structures, and centralization of scheduling tasks can lead to more rigor and objectivity. The result can be a reduction in labor costs and increased schedule quality (Maenhout & Vanhoucke, 2013b; Wright & Mahar, 2013).

Our research moves beyond an inter-departmental and centralized policy approach to the conduct of the Labor Planning and Allocation Cycle (LPAC) within a hospital to take a systemic, integrated approach to labor forecasting, scheduling, and staffing. Moreover, rather than optimizing the individual components of workload forecasting, scheduling, and allocation separately, this research considers all three components within their interdependent, dynamic, cyclical systemic nature and seeks to analyze the impact of the component interactions, organization structure, knowledge creation/sharing and human interactions involved within the work tasks. The research objective is to generate a new LPAC management model that leverages a systemic optimization of organizational structures, work tasks and human interactions based on patient flow to create improved outcomes. We use an Action Design Research (ADR) (Mullarkey & Hevner, 2015) method to the guided emergence of two main innovative artifacts – the Systemic LPAC Model and LPAC Model Outcome Metrics – that we evaluate and improve through interventions in situ with practitioners. We ask the research questions: what will a better future look like with a Systemic LFSS deployed in a typical hospital? How will a systems approach driven by patient flow and not limited by departmental structures lead to greater optimization of the balance between nursing labor, nursing satisfaction, and patient outcomes?

Motivation

The personnel capacity planning problem in any service-based industry involves strategic, tactical, and operational decision making concerning differing time horizons. Operational decision making can further be broken down into two subcategories including offline and online components. Offline decision making involves short-term decisions that take place in advance of operations while online decision making involves more reactive decisions adapting to real-time developments (Hulshof, Kortbeek, Boucherie, Hans, & Bakker, 2012; N. Kortbeek, Braaksma, Burger, Bakker, & Boucherie, 2015; Nikky Kortbeek, Braaksma, Smeenk, Bakker, & Boucherie, 2015).

The first decision level is strategic and involves the determination of the number of employees by department and skill set hired/retained to provide adequate resources for operations - capacity. Capacity calculations are typically performed annually as part of a budget process. The second decision level is tactical and includes two main components. The first component involves the determination of appropriate staffing levels including how many of each employee type is needed to meet the average workload demand - commonly referred to as the staffing plan for each department. The second component involves the recurring workload forecasting tasks performed before the creation of each labor schedule providing the guidelines for determining the actual number of each employee type scheduled per shift for the upcoming schedule period (next one, two or four weeks typically). These two sets of activities determine the number of shifts for each employee type needed for each shift/day of the schedule. The third decision level is operational involving both offline and online operational decisions. Offline decisions include matching specific staff names to specific shifts to complete the labor schedule. These activities are typically referred to as scheduling or rostering. Scheduling is proactive and typically involves a time horizon of four to six weeks. Online decisions include the adjustment of staff at the start of the shift to accommodate any last-minute changes in patient load, staff absence, or staff tardiness (Hulshof et al., 2012; N. Kortbeek et al., 2015). We chose to categorize these decision levels into *hiring*, developing staff *planning*, *forecasting* workload demand, labor *scheduling*, and *staffing* (Defraeye & Van Nieuwenhuyse, 2016; Thompson, 1993, 1995, 1998).

Within hospitals, the repeating tactical/operational cycle of labor planning and allocation consists of workload forecasting, labor scheduling, and labor staffing. The strategic task of determining the overall number of staff to hire/retain and the tactical task of developing departmental staffing plans are not part of the shorter-term, recurring LPAC cycle, but typically performed outside of the tactical/operational cycle (refer to Figure 2). Each of the activities within the tactical LPAC cycle plays a part in the overall process and can impact each

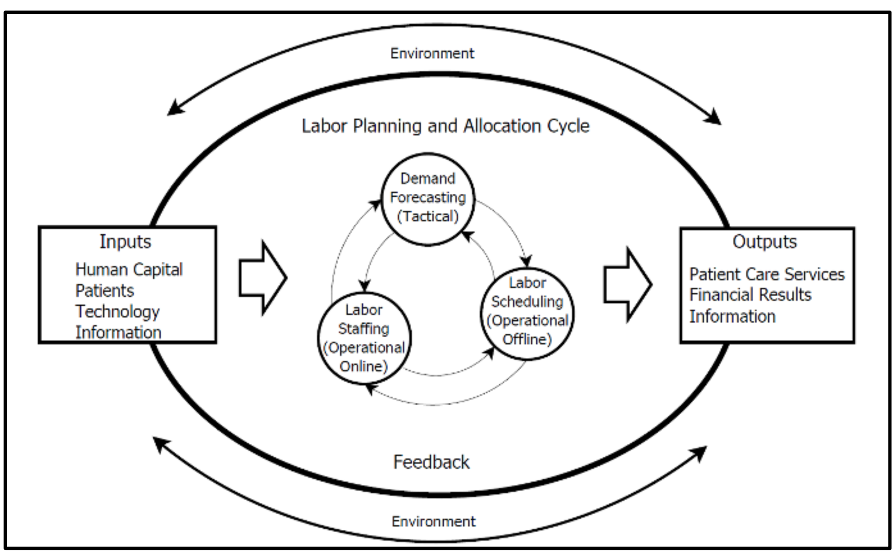


Figure 2 - Integrated Forecasting, Scheduling, and Staffing (LPAC Cycle)

downstream activity within the process (Nikky Kortbeek et al., 2015; Maenhout & Vanhoucke, 2013a). Interaction can be top-down imposing restrictions and constraints on downstream decisions or interaction can be bottom-up in the form of feedback to higher level decisions (Hulshof et al., 2012).

The two-way (Strategic Capacity vs. Tactical Staffing) interactional model of the LPAC cycle provides one layer of complexity. There are three additional layers of complexity that make the overall process extremely challenging in health care. The lowest most detailed layer includes the complex nature of quantifying workload within a hospital department. In particular, the patient needs matter in healthcare and are highly variable, and levels of care are critical to outcomes. Seemingly simple calculations of the optimal balance of nursing labor include the dynamic factors surrounding the given number of patients, their acuity (severity of illness), the variability of required services, and various other systemic factors like seasonality of demand and environmental conditions. Types and amounts of workload per patient can be highly variable and change dynamically shift-to-shift. They are notoriously difficult to accommodate and plan for during capacity, planning and forecasting activities. This complex nature of workload optimization – the right nurse at the right time for the right level of patient acuity – is impacted by the LPAC cycle constructed from a collection of departments interacting within several sub-systems or service lines (refer to Figure 3).

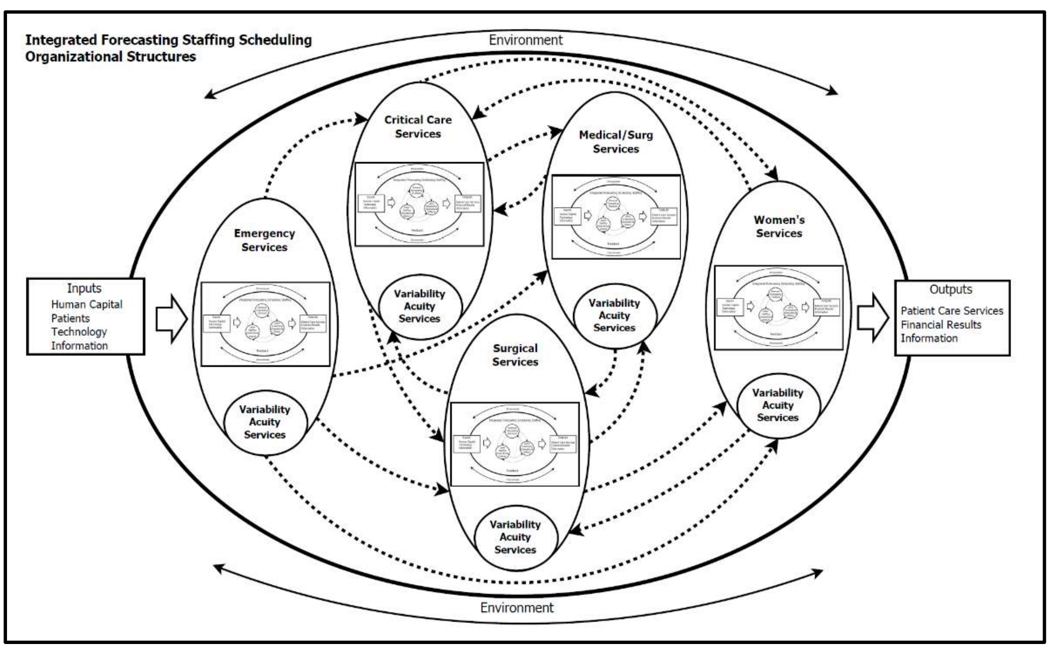


Figure 3 - The Complexity of Labor Planning and Allocation Within a Hospital

Inherently, the LPAC cycle in hospitals is complex, highly variable and dynamic. Unless we consider the system holistically, it is easy to sub-optimize in any given department – efficiently plan and employ the nursing staff, accommodate patient acuity, and generate optimal patient outcomes. And then, the very next shift, a change in any one of these variables can lead to an unsatisfactory result for nurse, patient, and hospital. LPAC successful outcomes are elusive for hospitals because they remain operationally siloed along departments and service lines. Therefore, even though the service lines and departments can be interdependent when it comes to patient flow – urgent care patients tend to move to specialty or general care departments, for example – nurse capacity, planning, and staffing isolate the caregivers in individual departments.

The problem confronted in this research is complex and not often studied at the organizational level of analysis. No generalized artifacts exist to address the structure of work tasks, human interactions, and leadership using the LPAC model to guide successful outcomes systemically for the hospital. While the challenges of labor-management outcomes have been well documented and recognized within academic and practitioner environments, few macro-level studies exist to study interactions, interdependency, patient flow, and integrated knowledge generation/sharing within a staffing cycle. Over the last decade, information technology and, to a lesser extent, process modeling have been applied to the problem, but there has been no significant impact on overall labor productivity within the industry. While the overall United States long-term average annual economic productivity gain is estimated to be approximately 1.1%, productivity gains achieved in the hospital sector have been closer to 0.4% (Cylus & Dickensheets, 2008).

Research Design

This research proposes that a gap exists between the desire to optimize LPAC outcomes and the actual results of existing LPAC instantiations in hospitals. Technology systems that automate existing tasks fail to account for human judgment, communication, interaction, and sense-making. Our project's objective is to document an understanding of the critical inputs to LPAC cycle components; knowledge generated and shared within each component and between each component, work and task structures that impact knowledge creation and sharing, and organizational structures that facilitate confirming and corrective interventions. We postulate that this understanding can subsequently be leveraged to re-shape the Systemic LPAC model as the framework for designing new work and organizational structures/systems to facilitate outcome improvement. The complexity of this objective requires a research method that goes beyond observation and explanation. Consequently, we chose a research method grounded in action research that involves an intervention with practitioners in situ to co-create and co-evaluate multiple iterations of a future state Systemic LPACe that optimizes the balance between nursing labor, cost, and patient outcomes.

We use the four-stage elaborated Action Design Science Research (eADR) model (Mullarkey & Hevner, 2015) to structure our research method. Embedding researchers within the practitioner environment provides the iterative opportunity to design and refine artifacts as used within the LPAC. Through the iterative use – define, build, evaluate, learn, reflect, re-define process – we find these artifacts become increasingly more relevant and contribute more toward the solution.

The first phase of the research was conducted in a midsize hospital in the southeastern United States. The hospital had sufficient data concerning labor cost, staff scheduling/allocation, staff dissatisfaction, and patient outcomes to recognize that a problem existed. In spite of large investments in technology over the years, the hospital routinely failed to meet its objectives in all of these categories. In partnership, researchers and hospital leadership outlined two overarching objectives: create the ability to quantify and measure outcomes of each successive phase of the LPAC artifact and investigate a structure of LPAC management to optimize outcomes. Fulfilling these objectives required an intervention over multiple iterative artifact design, build, and evaluation steps in situ. As a research-in-progress paper, we focus on the initial efforts of the ADR Problem Diagnosing (PD) stage for defining and framing the problem including determining outcome metrics for the various stages of the LPAC cycle.

LPAC Phase Performance Measurement

Within a system, each production subsystem's function is to process or transform inputs into outputs (Katz & Kahn, 1978). The majority of inputs are information and knowledge, while the outputs consist of various artifacts that are either fed into subsequent components or ultimately, result in the final state of staffing at the end of each shift. The transformation of inputs into outputs essentially is performed via the structure and workflow within the system. Our starting point for the analysis of structures and workflows was the determination of outcome metrics to assess performance for each LPAC stage.

The LPAC is essentially a service function that is a contributor to higher organizational service functions such as patient care. Measuring service quality is inherently more difficult than measuring product quality (Johnson et al., 2006). We recognize challenges in measuring the performance of the LPAC as a service-oriented function in line with previous research. However, to establish performance baselines as well as assess individual stage and overall system performance, a set of performance metrics is needed. The research objective of this phase was to test various potential LPAC stage outcome metrics through multiple iterations of real system data to determine a viable outcome measurement methodology. Our action research design started with the definition of metric acceptance criteria and the selection of potential metrics for each LPAC phase. We then executed three iterations of artifact define, build, evaluate, learn, reflect, and redefine steps with actual system data to refine the proposed artifact-metrics to meet the agreed upon acceptance criteria. Once completed, the research team was confident that the set of artifact-metrics would accomplish the stated goal of measuring each LPAC stage and provide relevant information for practitioners to take either corrective or confirming actions.

The effort began with a focused review of existing methodologies for measuring each LPAC stage (workload forecasting accuracy, scheduling, and staffing). Four criteria were developed to use in the evaluation of potential metrics suitability. The first criterion was that any evaluation of performance must have credibility with practitioners to be useful as a potential transformative artifact. Therefore, this acceptance criterion required that each metric be sufficiently easy to understand to be operationalizable in real-life situations. The measure must provide relevant information to quickly and accurately diagnose the state of any of the individual stages of the LPAC to elicit either a confirming or a correction action (Adoption Requirement). This criterion is critical from the aspect of practitioner usefulness. The second criterion required that each metric provide the ability for valid comparisons across multiple data series and across different time horizons or time series (Time Series Comparisons Requirement). The ability to compare outcomes from different time series was necessary for pre and post-implementation comparisons to evaluate overall model performance. The third criterion required that each metric provide the ability for valid comparisons across multiple data series from different hospital departments and different hospitals (Organizational Level Comparisons Requirement) to compare model performance across different organizational levels. The final criterion required that the metric be able to be “rolled-up” to aggregate departments into a group assessment (System Perspective Requirement). This requirement was critical to provide the ability to combine departments for measuring groups of departments as sub-systems (service lines) of the larger overall hospital system.

Workload Demand Forecasting Metric Selection

Workload demand forecasting is critical to personnel capacity decisions in all industries where throughput systems exist since typically inventory is the representation of workload. It is important, however, to recognize a key distinction with hospitals. Hospital inpatient (IP) units or nursing units can essentially be considered a complex throughput system including different types of patients (inventory) moving through the system via different patterns. Patients will have different arrival sources, different patient paths through the various departments based on patient conditions, illness, and acuity and different discharge destinations (Broyles, Cochran, & Montgomery, 2010). The understanding of workload demand makes up the foundation upon which labor schedules are built and therefore is a critical component contributing to the success of the overall labor planning process. An organization must understand the accuracy of demand forecasts to create labor schedules in support of the workload forecasted.

There is a large quantity of forecast accuracy measurement research that has consistently demonstrated that there is no one "best" measurement that works in all situations. Therefore, the best accuracy measurement is one that improves decision making and meets the needs of the resources using the forecast (Makridakis, 1993). Our research retained the same objective. We started with forecasting accuracy measurements that have been shown in prior research to be able to accommodate shift-based, nonstationary demand. These measurements were first analyzed to eliminate any potential measurement that could be shown to be invalid within a hospital workload forecasting situation. This analysis eliminated the three relative error measurements since research indicated that these measurements would be invalid within intermittent demand or zero error conditions causing a divide by zero error (Hyndman, 2006). Additionally, the Mean Absolute Scale Error measurement was eliminated due to its complexity.

Table 1 - Initial Workload Forecasting Accuracy Measures

Category	*Measurement*	*Abbreviation*
Scale dependent	Mean Average Error	F-MAE
	Mean Square Error	F-MSE
	Root Mean Square Error	F-RMSE
Percentage Error	Mean Absolute Percentage Error	F-MAPE
	Median Absolute Percentage Error	F-MdAPE

Table 1 details the initial set of demand forecast accuracy metrics. The metrics were then run through three iterations of the define, build, evaluate, learn, reflect, re-define process. The goal of each iteration was to evaluate the metrics against the five acceptance criteria to determine the suitability and refine the calculations based on conversation and learning.

Schedule Quality Metric Selection

Following the creation of an accurate workload forecast, work begins to create a labor schedule that assigns staff to slots within each shift to provide enough labor to accommodate the forecasted demand. Scheduling essentially represents two main objectives: providing enough labor to fill all of the necessary slots and providing the slotting in a manner that serves to maintain staff satisfaction. The first objective of scheduling, the right amount of labor contributes to labor cost control in two main ways. The minimization of under-scheduling will minimize the potential need for premium labor (i.e., overtime, contract labor, incentive pay) often used to provide last-minute labor to fill unfulfilled scheduled needs. In the opposite scenario, the minimization of over-scheduling will reduce the potential for too many staff on a given shift due to poor allocation decisions. Since the labor schedule is the starting point for all staffing decisions, a schedule mapped to an accurate demand forecast reduces the effort required during the staffing process.

Early nurse schedule quality research comes from a 1996 dissertation by Johan Oldenkamp who identified five main areas of schedule quality: completeness, optimality, proportionality, schedule healthiness, and continuity (Oldenkamp, 1996). Completeness refers to a schedule's ability to meet the quantitative demands for services in the unit (i.e., number of each staff member needed). Optimality represents the degree nursing expertise is equally distributed across the shifts. Proportionality refers to the degree of distribution of less desirable shifts across each staff member. Schedule Healthiness refers to the degree of unhealthy shift patterns present per staff member (Oldenkamp, 1996). Tarpey and Nelson revised the metrics to a simpler format with the intention of providing schedule quality metrics that were understandable by practitioners and more conducive to the application in real-world scenarios. The resulting metrics were: completeness, commitment, schedule healthiness, and preferences. Completeness refers to the percentage of overfilled or underfilled shifts in a given schedule indicating how well a schedule maps to the labor required for a specific demand forecast. Commitment refers to the percentage of employees who are scheduled to their full commitment level or a minimum number of hours expected to work (i.e., full allocation of hours). Schedule healthiness has the same definition as proposed by Oldenkamp referring to the percentage of healthy versus unhealthy shift patterns. This metric provides an indication of work hours that may lead to staff fatigue and ultimately stress and burnout. Preferences refer to the percentage of staff preferences honored in a given schedule (Tarpey & Nelson, 2009). After considerable review and discussion concerning the various schedule quality metrics, the team determined that the simpler set of metrics would be the logical starting point for this project (refer to Table 2).

Table 2 – Initial Schedule Quality Measures

Category	*Measurement*	*Dimension*
Total Completeness	Measures the prospective schedule effectiveness in meeting departmental workload demand	Schedule to Employee Needs
Professional	Same as above, specific to nursing	
Support	Same as above, specific to support staff	
Commitments	Measures the prospective schedule effectiveness in scheduling all resources to full available hours	Schedule to Department Needs
Healthiness	Measures the prospective schedule effectiveness in scheduling healthy shift patterns	Schedule to employee needs
Preferences	Measures the prospective schedule effectiveness in meeting employee desires/preferences	Schedule to employee needs

Allocation (Staffing)

Labor staffing, the operational online decision level of the personnel capacity planning problem, can best be characterized as "adjusting individual work assignments to account for daily fluctuations in the patient population, absenteeism, and emergencies" (Bard & Purnomo, 2005). It is at this stage where labor costs, staff satisfaction, and patient care are most at risk. Overstaffing provides unnecessary labor above and beyond the amount required to accommodate the existing workload resulting in unnecessary labor spend. Understaffing has the potential to impact labor cost, patient care quality, and staff satisfaction. The use of strategies such as overtime, contract labor, or the use of flexible pool labor to fill open shifts at the last minute typically results in higher labor cost due to the higher base cost of these resources. Running a department with labor shortages results in patient care risk since not enough labor is present to accommodate the existing workload.

Many methodologies have been explored to attempt to quantify nursing workloads. In their recent but separate works, MacPhee et al. and Holden et al. discuss the growing trend of using human factors frameworks to describe workloads from the perspective of unit-level, job-level, and task-level. These authors define the unit level as the number of staff and skill mix concerns in the unit, job level as perceptions of the amount of work to be accomplished, and task-level as individual cognitive requirements to complete a task (Holden et al., 2013; MacPhee, Dahinten, & Havaei, 2017). Our concern focuses on the unit-level measurement as the one most controlled by the labor planning and allocation cycle. When looking at staffing, productive hours by direct patient care staff have been shown to be an effective measurement of staffing to predict patient outcomes (Park, Blegen, Spetz, Chapman, & De Groot, 2015). This methodology has the advantage of being fairly easy to compute and comprised of readily available data. Other more complex measurements that include acuity and intensity include more subjective data and are not readily available at most hospitals unless a sophisticated patient classification or acuity system is in use. Our test hospital did not utilize an acuity or patient classification system, so this data was not available leading us to focus on direct patient care staff hours

During initial discussions, participants suggested productivity measures for measuring the success of staffing to the budget. The team eventually ruled out productivity measurements since the real-time staffing situation is concerned with measuring the performance of the staffing function resulting in having the right number of each skill set in the unit at the start of the shift. There was less concern with measuring the performance of the actual management of the department concerning time and task management since this was more of a departmental tactical function outside of the LPAC. The inclusion of fixed costs in addition to variable labor costs clouds the information provided by the measurements. While productivity measures may be satisfactory for financial reporting and benchmarking, they fall short of providing the value needed to assess the performance of a particular LPAC process participant's ability to correctly staff a department. Therefore, we focused on the initial staffing measurements detailed in Table 3 for our initial artifact design.

Table 3 – Initial Staffing Accuracy Measures

Category	*Measurement*	*Abbreviation*
Scale-Dependent	Mean Absolute Error (hours) - Difference (Miss) from staffing target	S-*h*MAE
Percent Error	Mean Absolute Percent Error (hours)	S-*h*MAPE

Metric Analysis – Artifact Iteration #1

The first iteration of measurement investigation focused on mapping each calculated metric to each of the first four acceptance criteria (the fifth criterion reviewed in a later iteration). The combined research-practitioner team had the latitude to determine if analysis of any particular metric should halt due to failure to meet any of the thresholds presented by the acceptance criteria. The group decided to halt at this point. The first iteration collected respective data for each metric from the first test 4-week schedule period beginning on 8/27/2017 for seven departments. The selected departments represented three service lines: A, B, and C (medical or surgical units), D and E (step-down units), and F and G (critical care units). The same time series of data was used to calculate the metrics (provided in Tables 4, 5, and 6).

Table 4 – Workload Forecasting Metrics Iteration #1 Results

Department	F-MAE (Patients)	F-MSE	F-RMSE	F-MAPE (%)	F-MdAPE (%)
A (Med/Surg)	3.39	21.61	4.65	14.68%	10.88%
B (Med/Surg)	4.07	31.64	5.63	15.57%	13.04%
C (Med/Surg)	2.43	7.07	2.66	21.00%	21.43%
D (Step Down)	3.39	17.11	2.66	12.33%	11.11%
E (Step Down)	2.14	6.57	2.56	8.57%	8.00%
F (Critical Care)	1.71	4.43	2.10	13.69%	12.50%
G (Critical Care)	5.46	43.54	6.60	23.63%	23.81%

Table 5 – Schedule Quality Metrics Iteration #1 Results

Department	Completeness	Completeness (Professional)	Completeness (Support)	Commitment	Healthiness	Preferences
A (Med/Surg)	85.12%	86.16%	61.61%	100.00%	96.58%	91.86%
B (Med/Surg)	74.11%	79.91%	58.93%	100.00%	94.85%	89.74%
C (Med/Surg)	92.41%	88.10%	83.93%	100.00%	95.63%	82.50%
D (Step Down)	92.86%	91.37%	83.57%	100.00%	93.65%	87.82%
E (Step Down)	79.51%	61.48%	67.14%	100.00%	94.93%	87.42%
F (Critical Care)	79.62%	79.69%	78.57%	100.00%	93.46%	88.02%
G (Critical Care)	90.41%	89.88%	100.00	100.00%	90.05%	87.77%

Table 6 – Staffing Metrics Iteration #1 Results

Department	S-hMAE (Hours)	S-hMAPE (Hours)
A (Med/Surg)	10.18	12.38%
B (Med/Surg)	11.39	12.69%
C (Med/Surg)	3.91	7.85%
D (Step Down)	11.36	9.98%
E (Step Down)	13.76	11.51%
F (Critical Care)	8.57	8.37%
G (Critical Care)	17.12	11.47%

Since the test metrics were all selected based on available data, the first acceptance criterion was considered satisfied for each metric.

Workload demand forecasting analysis considered the following measurements: F-MAE, F-MSE, F-RMSE, F-MAPE, and F-MdAPE. Upon analysis of the first acceptance criterion, the review team immediately identified an issue. The metrics, as calculated, provided data regarding the accuracy of the number of patients forecasted but no information about the amount of labor needed to be scheduled to cover the anticipated workload demand. The metrics did not provide relevant information needed to feed into the LPAC scheduling phase and therefore failed this test requiring modification to provide relevant information for scheduling. Additional conversations resulted in an agreement that the Mean Square Error (F-MSE), Root Mean Square Error (F-RMSE), and the Median Absolute Percentage Error (F-MdAPE) were not intuitive enough or easily explained to nursing department leaders. Therefore, the team eliminated F-MSE, F-RMSE, and F-MdAPE for the next iteration. Given the first test failure, the team elected not to proceed to test the metrics against the remaining acceptance criteria and refine the forecast accuracy metrics in the next iteration.

Schedule data consisted of data collected from the scheduling system at the point schedules were completed and communicated to employees, which was effectively two weeks before the beginning schedule start date. Analysis and evaluation of the metrics against the acceptance criteria led to positive conclusions. Discussions with nursing department leaders resulted in an agreement that the three completeness metrics represented the quality of the schedules' "fit" to the labor forecasted. There was a clear understanding concerning the percent of shifts accurately filled in the schedule based on the number of shifts needed to accommodate the forecasted demand. For example, looking at department A, the schedule includes labor matched to 85.12% of the anticipated demand needed for all staff. Therefore approximately 15% of the schedule is either underfilled or overfilled compared to the forecasted need. The same measurement is broken down into nursing and support, indicating 86.16% of the nurse schedule mapped to forecasted demand and 61.61% of the support staff schedule mapped to forecasted demand. It was clear to nursing leaders that the support staff schedule was not as good as the nurse staff schedule confirming the ability of the completeness metrics to be used for comparisons within a department. Additional conversations with nursing leaders included discussions comparing units to each other. The metrics indicated that the department D nurse schedule was well mapped compared to the other units. Nursing leaders investigated within the scheduling system and found that this schedule did have fewer shifts left uncovered. With this information, the review team concluded that the schedule quality completeness metrics met the first four acceptance criteria.

Similar conversations with nursing leadership resulted in equal clarity around the commitment, healthiness, and preference metrics. They understood the commitment metric value of 100.00% indicated that every staff member on the schedule was scheduled the full allotment of hours available. The healthiness metric adequately indicated the percentage of unhealthy shift patterns in the schedule. Finally, the preferences metric indicated the percentage of staff scheduling preferences honored in the schedule. Nursing leaders discussed, at length, the results of the metrics focusing on how well their schedules mapped to the forecast, how the staff was scheduled to their commitments, the number of unhealthy shift patterns and how staff satisfaction was either being supported or possibly not supported by the honoring of preferences. The rich content of these conversations led us to conclude that the all of the proposed schedule quality metrics provided useful information about the quality of schedules meeting the first four acceptance criteria.

Staffing metric discussions with the facility and nursing department leaders resulted in some disagreement. The departmental leaders supported the two proposed metrics based on hours, while the facility leadership, specifically the CFO, made the case to convert the metrics to Full-Time Equivalent (FTE) measurements, which is a more common unit of measurement at the facility level. Each camp made convincing cases supporting the information conveyed by each type of measurement and also agreed that an indication of shortage versus overage was needed. In evaluating the metrics against the Time Series Comparison and Organizational Level Comparison criteria, there was some hesitancy. At issue was whether or not the metrics accurately conveyed how one department was performing versus another department. For example, did department D perform equally well as department B by having an S-*h*MAE value of 11.36 hours off target as compared to S-*h*MAE of 11.39 hours off target? The larger hour shortage in department G is a more significant concern where the patient to nurse ratio is 2:1 versus department B, where the patient to nurse ratio is 5:1. Each hour missed in the department G is more critical with the lower patient to nurse ratio. The same concern involved the S-*h*MAPE measurement. When comparing S-*h*MAPE values, department D appears to have performed better than four other departments, although the team did not feel this was the case given the patient to nurse ratios involved. The team concluded that the staffing metrics failed the acceptance criteria tests and needed to be refined in the next iteration to account for the patient to nurse ratio. Table 7 details Iteration #1 metric acceptance results.

Table 7 – Iteration #1 Metric Acceptance Results

Category	Metric	Adoption	Time Series Comparison	Organizational Level Comparison	Systems Perspective	Conclusion Next Iteration
Workload Forecast Accuracy	*F-MAE*	Passed	Not Tested	Not Tested	Not	**Include/Modify**
Workload Forecast Accuracy	*F-MSE*	Passed	Not Tested	Not Tested	Not	**Include/Modify**
Workload Forecast Accuracy	*F-RMSE*	Failed	Not Tested	Not Tested	Not	Drop
Workload Forecast Accuracy	*F-MAPE*	Failed	Not Tested	Not Tested	Not	Drop
Workload Forecast Accuracy	*F-MdAPE*	Failed	Not Tested	Not Tested	Not	Drop
Schedule Quality	*Completeness*	Passed	Passed	Passed	Not	**Include**
Schedule Quality	*Professional*	Passed	Passed	Passed	Not	**Include**
Schedule Quality	*Support*	Passed	Passed	Passed	Not	**Include**

Table 7 (Continued)

Category	Metric	Adoption	Time Series Comparison	Organizational Level Comparison	Systems Perspective	Conclusion Next Iteration
Schedule Quality	*Commitments*	Passed	Passed	Passed	Not	**Include**
Schedule Quality	*Healthiness*	Passed	Passed	Passed	Not	**Include**
Schedule Quality	*Preferences*	Passed	Passed	Passed	Not	**Include**
Staffing	*S-hMAE*	Passed	Failed	Failed	Not	**Include/Modify**
Staffing	*S-hMAPE*	Passed	Failed	Failed	Not	**Include/Modify**

Metric Analysis – Artifact Iteration #2

The second artifact development iteration began with design sessions to incorporate learning from the first iteration into new versions of the forecast accuracy and staffing metrics. The schedule quality metrics were not modified but evaluated through another iteration of data.

The design work for workload demand forecasting resulted in changes to the base of the metric calculations to accommodate the focus shift on labor forecasted versus patients forecasted. Three new metric options were determined to be feasible: Full-Time Equivalents (FTEs), Labor Hours, and Staffing Grid Labor Bins (refer to Table 8).

The staffing grid bin basis requires some explanation (refer to Figure 4 and Figure 5 for examples). Each forecasted or actual patient value is used to determine the number of hours, FTEs or staffing grid bin based on the department staffing matrix (plan). According to the measurement methodology converting the patient values to FTEs, hours, or staffing grid bins, the three modified test measurements were delineated as: F-*f*MAE (Mean Absolute Error FTEs), F-*h*MAE (Mean Absolute Error Hours), F-*Sb*MAE (Mean Absolute Error Staffing Bins), F-*f*MAE (Mean Absolute Percentage Error FTEs), F-*h*MAPE (Mean Absolute Percentage Error Hours), F-*Sb*MAPE (Mean Absolute Percentage Error Staffing Bins).

Table 8 – Workload Forecasting Accuracy Measures (Iteration #2)

Category	*Measurement*	*Abbreviation*
Scale dependent	Mean Average Error (FTE)	F-*f*MAE
	Mean Average Error (Hours)	F-*h*MAE
	Mean Average Error (Staffing Grid Bins)	F-*Sb*MAE
Percentage Error	Mean Absolute Percentage Error (FTE)	F-*f*MAPE
	Mean Absolute Percentage Error (Hours)	F-*h*MAPE
	Mean Absolute Percentage Error (Staffing Grid Bins)	F-*Sb*MAPE

Forecast Hours, FTE, and Staffing Grid Bins Determination

Date	Shift (12-hours each)	Forecast Patients	Forecasted FTE	Forecasted Hours	Forecasted Staffing Bin	Actual Patients	Actual FTE	Actual Hours	Actual Staffing Bin
1/1/2017	Day	21	8	(8*12)=96	6				
1/1/2017	Night	21	8	(8*12)=96	6				
1/2/2017	Day	19	6	(6*12)=72	5				
1/2/2017	Night	19	6	(6*12)=72	5				

Day Shift Grid (12 Hours)

Patient Min	Patient Max	RN	NA	Clerks	Grid Bin #
1	7	2	0	0	1
8	9	2	1	0	2
10	15	3	1	0	3
16	18	4	1	0	4
19	20	4	2	0	5
21	25	5	2	1	6
26	26	5	3	1	7
27	31	6	3	1	8

Night Shift Grid (12 Hours)

Patient Min	Patient Max	RN	NA	Clerks	Grid Bin #
1	7	2	0	0	1
8	9	2	1	0	2
10	15	3	1	0	3
16	18	4	1	0	4
19	20	4	2	0	5
21	25	5	2	1	6
26	26	5	3	1	7
27	31	6	3	1	8

Forecast FTE = aggregate number of resources from each skill mix in the staffing grid row corresponding to the number of forecast patients
Forecast Hours = Forecast FTE * number of hours in shift (12 for day and night shift in our case)
Forecast Bin = Grid Bin # in the staffing grid row corresponding to the number of the forecast patients

Figure 4 - Forecast Patients, FTE, and Bin Determination

Actual Hours, FTE, and Staffing Grid Bins Determination

Date	Shift (12-hours each)	Forecast Patients	Forecasted FTE	Forecasted Hours	Forecasted Staffing Bin	Actual Patients	Actual FTE	Actual Hours	Actual Staffing Bin
1/1/2017	Day	21	8	(8*12)=96	6	16	5	(5*12)=60	4
1/1/2017	Night	21	8	(8*12)=96	6	15	5	(5*12)=60	4
1/2/2017	Day	19	6	(6*12)=72	5	12	4	(4*12)=48	3
1/2/2017	Night	19	6	(6*12)=72	5	12	4	(4*12)=48	3

Day Shift Grid (12 Hours)

Patient Min	Patient Max	RN	NA	Clerks	Grid Bin #
1	7	2	0	0	1
8	9	2	1	0	2
10	15	3	1	0	3
16	18	4	1	0	4
19	20	4	2	0	5
21	25	5	2	1	6
26	26	5	3	1	7
27	31	6	3	1	8

Night Shift Grid (12 Hours)

Patient Min	Patient Max	RN	NA	Clerks	Grid Bin #
1	7	2	0	0	1
8	9	2	1	0	2
10	15	3	1	0	3
16	18	4	1	0	4
19	20	4	2	0	5
21	25	5	2	1	6
26	26	5	3	1	7
27	31	6	3	1	8

Actual FTE = aggregate number of resources from each skill mix in the staffing grid row corresponding to the number of actual patients
Actual Hours = Actual FTE * number of hours in shift (12 for day and night shift in our case)
Actual Bin = Grid Bin # in the staffing grid row corresponding to the number of the actual patients

Figure 5 - Actual Patients, FTE, and Bin Determination

The new workload demand forecast metrics were calculated using the next 4-week data set beginning on 9/24/2017 (results detailed in Table 9) and reviewed against the acceptance criteria. Now that the metrics reflected labor rather than patients, the conversations shifted to learning from the first iteration concerning the patient to nurse ratio. The review team agreed that the staffing grid bin measurements provided the best indication of the accuracy of a forecast across the first four acceptance criteria since this methodology adjusted for the patient to nurse ratio. While all metrics provided some meaningful information when tested across the nursing department leaders, the F-*Sb*MAE and F-*Sb*MAPE measurements provided the most value when attempting to compare one department versus another. The F-hMAE, F-*f*MAE, F-*h*MAPE, and F-*f*MAPE measurements were all scale dependent on the patient to nurse ratio precluding the ability to determine whether one department was performing better than another.

With this information, the review team concluded that the F-*h*MAE, F-*f*MAE, F-*h*MAPE, and F-*f*MAPE measurements failed the Time Series and Organizational Level Comparison criteria. The F-*Sb*MAE and F-*Sb*MAPE measurements passed the first four of the acceptance criteria and retained for the third iteration.

Table 9 – Workload Forecasting Metrics Iteration #2 Results

Department	F-*h*MAE (Hours)	F-*f*MAE (FTE)	F-*Sb*MAE (Staffing Bins)	F-*h*MAPE (%)	F-*f*MAPE (%)	F-*Sb*MAPE (%)
A (Med/Surg)	5.75	1.44	0.79	23.15%	23.15%	11.87%
B (Med/Surg)	4.27	1.07	0.68	16.06%	16.06%	9.40%
C (Med/Surg)	1.95	0.49	0.64	14.21%	14.21%	18.10%
D (Step Down)	5.88	1.45	1.46	16.05%	16.05%	14.56%
E (Step Down)	6.79	1.70	0.89	20.66%	20.66%	11.76%
F (Critical Care)	3.14	0.79	0.64	9.91%	9.91%	7.47%
G (Critical Care)	6.55	1.64	1.00	16.31%	16.31%	9.94%

The schedule quality metric analysis generated discussion around the continued acceptance of the proposed metrics. Nursing leadership conversations quickly centered on which department was doing better compared to the other departments. It was clear that department E was having trouble scheduling, which was confirmed by the department director explaining a large number of vacancies (nurses) in the unit and other potentially mitigating factors. The under-scheduling of nurses, due to the number of nurse vacancies was represented in the low Professional Completeness metric of 51.79% (refer to Table 10). We felt confident that these were exactly the type of conversations desired to be generated by these metrics. The team determined that the metrics again met all of the first four acceptance criteria, but also decided that some form of information needed regarding whether the underlying issue was over-scheduling or under-scheduling. We decided to address this need during the third iteration.

The same thought process concerning the removal of the patient to nurse ratio dependency impacted the staffing metric decisions in the second iteration. The original S-*h*MAE and S-*h*MAPE metrics were converted to FTE and staffing grid bin metrics using the same logic as was applied to the workload demand forecasting metrics (refer to Table 11 for the metrics).

Table 10 – Schedule Quality Metrics Iteration #2 Results

Department	Completeness	Completeness (Professional)	Completeness (Support)	Commitment	Healthiness	Preferences
1E (Med/Surg)	76.64%	71.16%	66.20%	100.00%	99.01%	88.40%
1W (Med/Surg)	71.43%	78.13%	56.25%	100.00%	95.91%	90.37%
2E (Med/Surg)	94.20%	88.10%	87.50%	100.00%	96.84%	85.98%
PCU (Step Down)	91.60%	91.67%	84.29%	100.00%	95.84%	83.15%
SICU (Step Down)	60.79%	51.79%	82.32%	100.00%	94.16%	88.75%
CCU (Critical Care)	79.62%	79.67%	78.57%	100.00%	91.01%	91.05%
ICU (Critical Care)	88.53%	87.10%	85.71%	100.00%	91.03%	94.66%

Table 11 – Staffing Accuracy Measures (Iteration #2)

Category	*Measurement*	*Abbreviation*
Scale-Dependent	Mean Absolute Error (hours) - Difference (Miss) from staffing target	S-*h*MAE
	Mean Absolute Error (FTE) - Difference (Miss) from staffing target	S-*f*MAE
	Mean Absolute Error (Staffing Bin) - Difference (Miss) from staffing target	S-*Sb*MAE
Percent Error	Mean Absolute Percent Error (hours)	S-*h*MAPE
	Mean Absolute Percent Error (FTE)	S-*f*MAPE
	Mean Absolute Percent Error (Staffing Bin)	S-*Sb*MAPE

Discussions on iteration #2 of the staffing metric results were long. In the end, the team determined that the S-*f*MAE and S-*Sb*MAE metrics provided the most relevant and useful information concerning staffing (refer to Table 12 for results). The S-*Sb*MAE metric provided the average number of staffing grid bins the department operated away from the target. Comparing departments A and C, the metrics show that department C ran closer to target by only being 0.29 bins away from the target on the average. Since the metric indicates staffing bins and the bins are already adjusted for the patient to nurse ratio, the metric provided a scale-independent method for comparing departments. In the long run, a department wants to operate within the correct staffing grid bin indicating that the department is following its staffing plan. The closer S-*Sb*MAE is to 0.00, the closer the department is running to their target. Nursing departmental leaders were able to understand the metric and compare one unit to another in performance. The S-*f*MAE metric provided the same information, but discussion still hung on the inability to determine if the average FTE miss was acceptable or not depending on the patient to nurse ratio. The review team concluded that the S-*f*MAPE and S-*Sb*MAPE measurements did not provide intuitive information for the department leaders and therefore failed the Adoption criterion. Therefore, S-*f*MAE and S-*Sb*MAE carried into the third iteration since they passed all of the first four acceptance criteria. Table 13 provides a summary of acceptance conclusions.

Table 12 – Staffing Metrics Iteration #2 Results

Department	S-*h*MAE (Hours)	S-*h*MAPE (Hours)	S-*f*MAE (FTE)	S-*f*MAPE (FTE)	S-*Sb*MAE (Staffing Bin)	S-*Sb*MAPE (Staffing Bin)
A (Med/Surg)	8.98	9.86%	0.75	9.86%	0.82	13.66%
B (Med/Surg)	8.27	8.94%	0.69	8.94%	0.77	12.27%
C (Med/Surg)	3.56	7.26%	0.30	7.26%	0.29	8.04%
D (Step Down)	9.54	8.33%	0.79	8.33%	0.86	9.54%
E (Step Down)	11.05	10.36%	0.92	10.36%	0.96	13.08%
F (Critical Care)	8.54	7.80%	0.71	7.80%	0.75	11.26%
G (Critical Care)	6.98	5.90%	0.58	5.90%	0.36	4.08%

Table 13 – Iteration #2 Metric Acceptance Results

Category	Metric	Data Available	Adoption	Time Series Comparison	Organizational Level Comparison	Systems Perspective	Conclusion Next Iteration
Workload Forecast Accuracy	*F-hMAE*	Passed	Passed	Failed	Failed	Not Tested	Drop
Workload Forecast Accuracy	*F-fMAE*	Passed	Passed	Failed	Failed	Not Tested	**Include**
Workload Forecast Accuracy	*F-SbMAE*	Passed	Passed	Passed	Passed	Not Tested	**Include**
Workload Forecast Accuracy	*F-hMAPE*	Passed	Passed	Failed	Failed	Not Tested	Drop
Workload Forecast Accuracy	*F-fMAPE*	Passed	Passed	Failed	Failed	Not Tested	Drop
Workload Forecast Accuracy	*F-SbMAPE*	Passed	Passed	Passed	Passed	Not Tested	Drop
Schedule Quality	*Completeness*	Passed	Passed	Passed	Passed	Not Tested	**Include**
Schedule Quality	*Professional*	Passed	Passed	Passed	Passed	Not Tested	**Include**
Schedule Quality	*Support*	Passed	Passed	Passed	Passed	Not Tested	**Include**
Schedule Quality	*Commitments*	Passed	Passed	Passed	Passed	Not Tested	**Include**
Schedule Quality	*Healthiness*	Passed	Passed	Passed	Passed	Not Tested	**Include**
Schedule Quality	*Preferences*	Passed	Passed	Passed	Passed	Not Tested	**Include**
Staffing	*S-hMAE*	Passed	Failed	Failed	Failed	Not Tested	Drop
Staffing	*S-hMAPE*	Passed	Failed	Failed	Failed	Not Tested	Drop
Staffing	*S-fMAE*	Passed	Passed	Passed	Passed	Not Tested	**Include**
Staffing	*S-fMAPE*	Passed	Failed	Failed	Failed	Not Tested	Drop
Staffing	*S-SbMAE*	Passed	Passed	Passed	Passed	Not Tested	**Include**
Staffing	*S-SbMAPE*	Passed	Failed	Failed	Failed	Not Tested	Drop

Metric Analysis – Artifact Iteration #3

The third iteration included the last assessment of the remaining metrics as well as the aggregation of department detail to determine the metrics' ability to measure the performance of systems of departments. The team felt positive about the metrics left in contention with the only exception being the remaining concern about the indication of shortage and overages. The team determined that the only two feasible options were to either separate each metric into an overage and underage metric or retain the metrics and provide a visual to indicate performance. The split metric option was determined to be potentially too confusing. Therefore, the team worked to create the visual indications needed.

Workload demand forecasting work focused on testing and evaluation of the remaining metrics F-*Sb*MAE and F-*Sb*MAPE and the creation of visuals. Table 14 details calculated results from the third iteration. Final discussions between the team and the departmental nursing leaders concluded that the F-*Sb*MAE metric provided the most intuitive information and the ability to compare across different schedule periods and different organizational levels. When considering visual representation, the team determined that a good target for forecasting would be a forecast within ±1 staffing grid bin of the actual bin number. With this information, the team developed a control chart for workload demand forecasting to provide the overage and underage information. The two graphs in Figure 6 provide information of when the workload forecast varies outside the control threshold. Based on this data, it is easy to understand that the department E forecast was a more accurate forecast than the department G forecast.

Table 14 – Workload Forecasting Metric Iteration #3 Results

Department	F-*Sb*MAE	F-*Sb*MAPE
A (Med/Surg)	1.14	16.45%
B (Med/Surg)	1.18	15.20%
C (Med/Surg)	0.36	10.48%
D (Step Down)	1.50	14.76%
E (Step Down)	0.29	3.51%
F (Critical Care)	1.11	14.18%
G (Critical Care)	1.71	29.77%

The third design session for schedule quality provided no calculation changes from the second iteration. The team focused on developing visual representations of the metrics to show where schedules were over or under allocated and analyzing the metrics' ability to be aggregated into groups of departments. Table 15 details schedule quality metric values from the third iteration.

Table 15 – Schedule Quality Metric Iteration #3 Results

Department	Completeness	Completeness (Professional)	Completeness (Support)	Commitment	Healthiness	Preferences
A (Med/Surg)	63.81%	58.57%	74.29%	100.00%	93.75%	87.10%
B (Med/Surg)	77.14%	75.71%	75.71%	100.00%	94.93%	87.20%
C (Med/Surg)	83.04%	74.04%	91.07%	100.00%	96.89%	76.30%
D (Step Down)	87.82%	90.18%	76.43%	100.00%	94.90%	88.90%
E (Step Down)	71.24%	65.56%	87.14%	100.00%	93.66%	88.80%
F (Critical Care)	83.82%	84.38%	78.00%	100.00%	92.13%	91.30%
G (Critical Care)	80.28%	80.03%	85.71%	100.00%	92.06%	88.60%

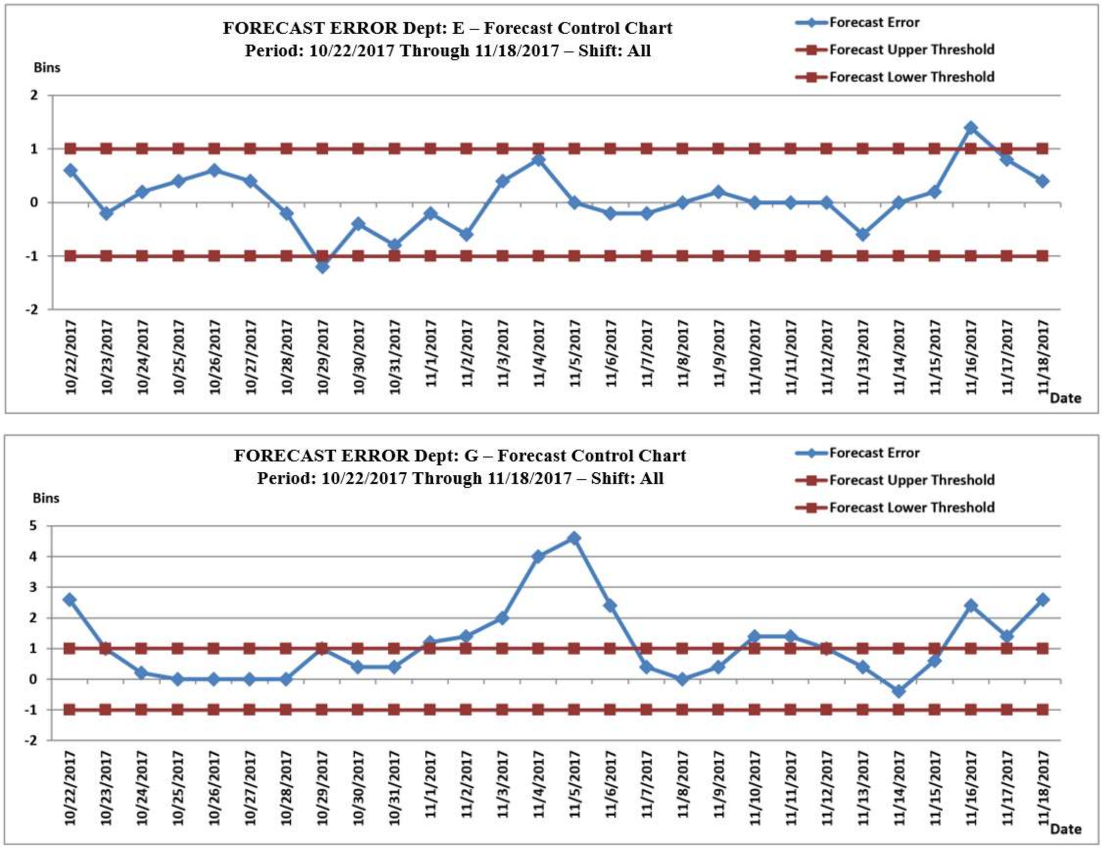


Figure 6 - Workload Forecasting Control Charts Iteration #3 Results

The conversations with the nursing department directors in the third iteration centered on improvements and declines from the prior two iterations. These conversations were exciting to listen to as they supported the intent of the metrics to measure performance across different time periods and organizational levels. The conversations were productive as they delved into more discrete reasons for poor or good scheduling. The discrete reasons, however, confirmed the need for practitioners to have a visual concerning the over or under-scheduling to consider corrective actions. These discussions led to the development of the graphs presented in Figure 7. Nursing leaders agreed these graphs were useful indicators providing detail of where the schedule completeness issues existed. They also provided information useful in time series and organizational level comparisons.

The third design session for staffing focused on S-*f*MAEf and S-*Sb*MAE. The team concluded that the metric that provided the most relevant, easy to understand scale independent information was the S-*Sb*MAE metric due to its resistance to the patient to nurse ratio issues. Along with a developed visual, the team and departmental nursing leaders felt that staffing per unit could be measured and compared across schedule periods and departments (refer to Table 16 and Figure 8).

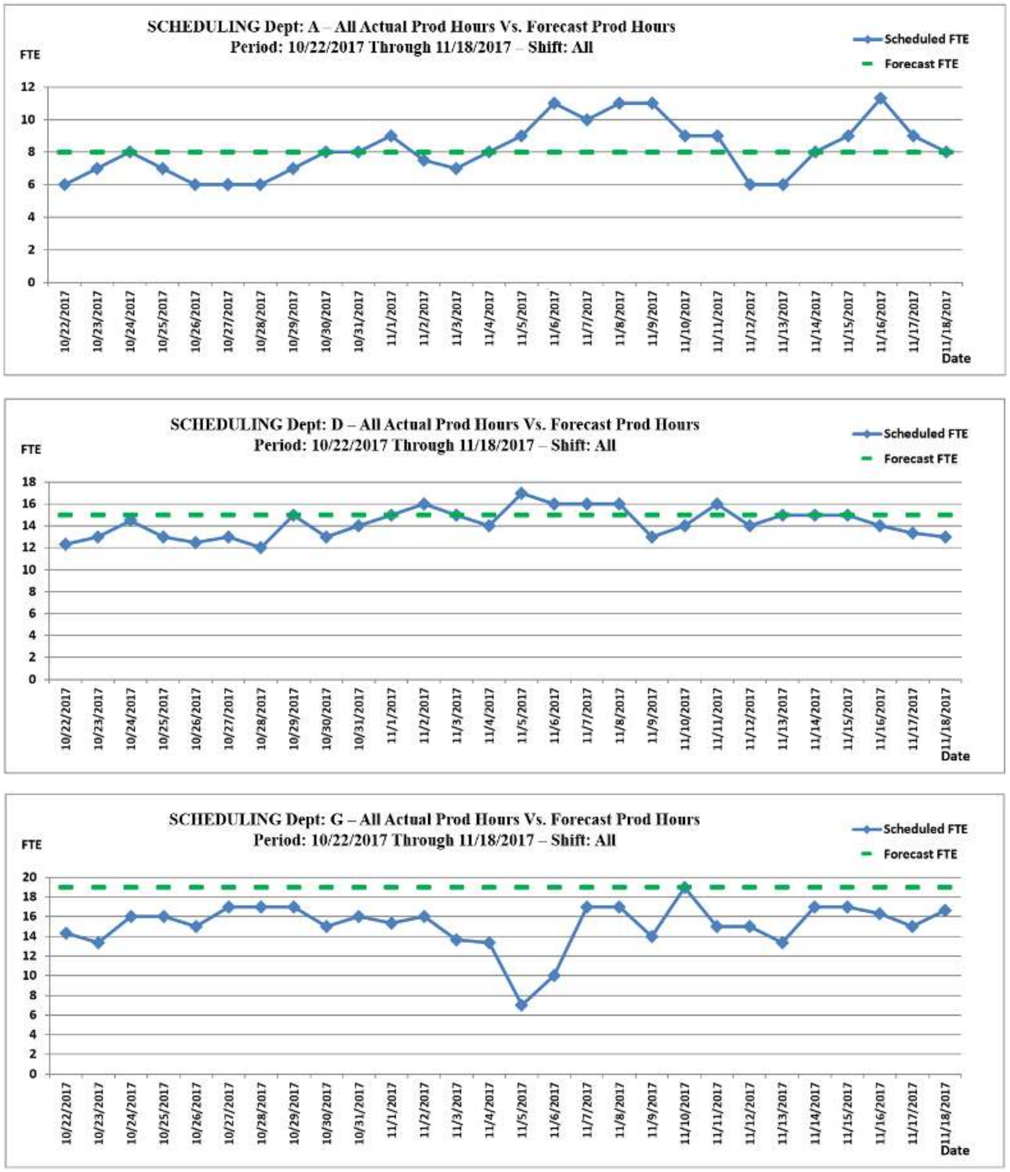
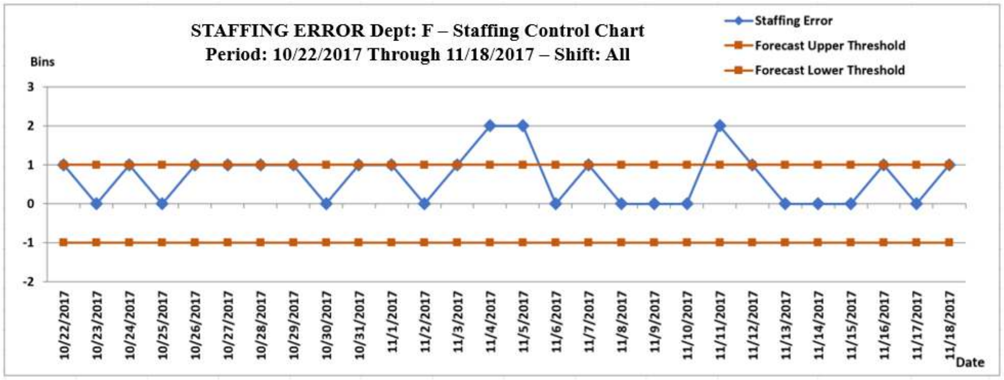


Figure 7 - Schedule Quality Completeness Graphs

Table 16 – Staffing Metric Iteration #3 Results

Department	S-*f*MAE (FTE)	S-*Sb*MAE (Staffing Bin)
A (Med/Surg)	0.12	0.07
B (Med/Surg)	0.55	0.61
C (Med/Surg)	0.12	0.14
D (Step Down)	0.49	0.57
E (Step Down)	0.41	0.50
F (Critical Care)	0.67	0.68
G (Critical Care)	0.45	0.39



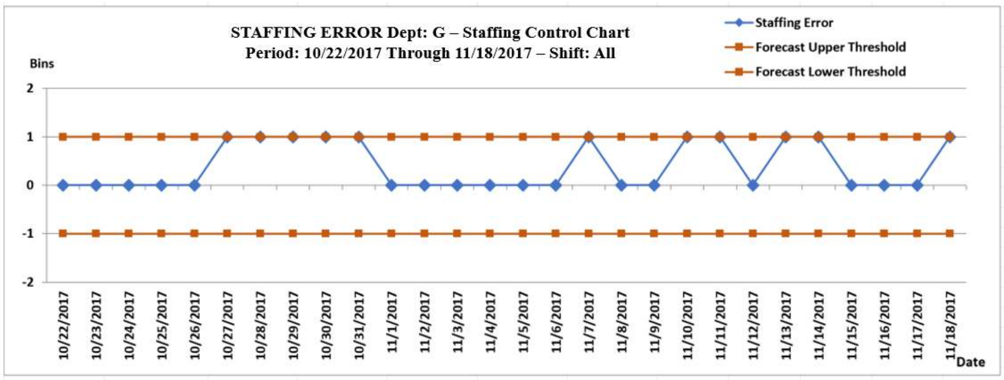


Figure 8 - Staffing Control Charts Iteration #3 Results

The visual consisted of standard control charts indicating data points outside of the threshold of plus one or minus one staffing grid bin. Each data point outside of the control threshold can be considered a point of concern since the staffing is well off target with an indication of either over or understaffing. Nursing leaders responded well to these charts as a mechanism to easily assess staffing performance on any given day.

System Based Measurement

During the third iteration, work was done to aggregate each of the metrics to measure performance across service lines (e.g., medical/surgical, step down, and critical care) to determine adherence to the final acceptance criterion. Each of the metrics was re-calculated by aggregating the individual departments according to their service line (results provided in Table 17). The subsequent results proved useful in viewing the performance of each ISFF phase within each service line. Nursing leaders agreed that these measurements along with the visuals were useful in determining if a particular service line had the correct amount of labor either forecasted, scheduled, or staffed. This concept is important when considering the departments as sub-systems of departments within the overall hospital system. If one department within a service line is over-scheduled and another is under-scheduled, there is value in knowing if the amount of staff across the entire service line will cover the anticipated need. Staff can subsequently move where the need exists.

The data in Table 17 provide a reasonable mechanism for comparing service lines. Visuals included in Figures 9 through 11 provided information concerning overages and underages. Nursing leadership saw clear opportunities when viewing the metrics at the service line level. Forecasting and scheduling were concerns due to their indication of being well off target. The nursing leaders concluded that poor results from the planning phases of the LPAC likely impacted. They agreed that the final set of metrics provided relevant, previously unknown, information measuring the performance of each LPAC phase for the individual departments as well as service lines and agreed that the final set of metrics met the five acceptance criteria (refer to Table 18).

Table 17 – Allocation Iteration #3 Results (Service Line)

Service Line	Forecast Accuracy	Schedule Quality						Staffing
	F-*Sb*MAE (Staffing Bins)	Completeness	Completeness (Professional)	Completeness (Support)	Commitment	Healthiness	Preferences	F-*Sb*MAE (Staffing Bin)
Med/Surg	0.738	73.12%	68.82%	77.68%	100.00%	94.97%	84.73%	0.274
Step Down	0.893	79.07%	76.92%	81.79%	100.00%	94.31%	88.84%	0.536
Critical Care	1.411	81.79%	81.86%	80.36%	100.00%	92.09%	89.80%	0.536

Table 18 – Final Metric Acceptance Results

Category	Metric	Adoption	Time Series Comparison	Organizational Level Comparison	Systems Perspective	Final Conclusion
Workload Forecast Accuracy	***F-SbMAE***	Passed	Passed	Passed	Passed	**Accept**
Workload Forecast Accuracy	*F-fMAE*	Passed	Failed	Failed	Passed	Drop
Schedule Quality	***Completeness***	Passed	Passed	Passed	Passed	**Accept**
Schedule Quality	***Professional***	Passed	Passed	Passed	Passed	**Accept**
Schedule Quality	***Support***	Passed	Passed	Passed	Passed	**Accept**
Schedule Quality	***Commitments***	Passed	Passed	Passed	Passed	**Accept**
Schedule Quality	***Healthiness***	Passed	Passed	Passed	Passed	**Accept**
Schedule Quality	***Preferences***	Passed	Passed	Passed	Passed	**Accept**
Staffing	*S-hMAE*	Failed	Failed	Failed	Not Tested	Drop
Staffing	*S-hMAPE*	Failed	Failed	Failed	Not Tested	Drop
Staffing	*S-fMAE*	Passed	Passed	Passed	Passed	**Accept**
Staffing	*S-fMAPE*	Failed	Failed	Failed	Not Tested	Drop
Staffing	*S-SbMAE*	Passed	Passed	Passed	Passed	**Accept**
Staffing	*S-SbMAPE*	Failed	Failed	Failed	Passed	Drop

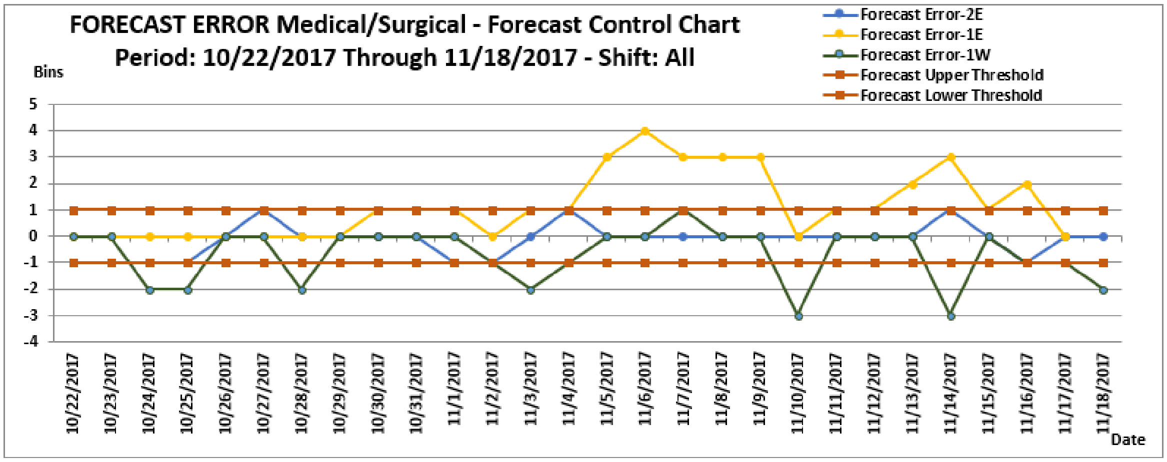
Conclusions, Limitations, and Future Directions

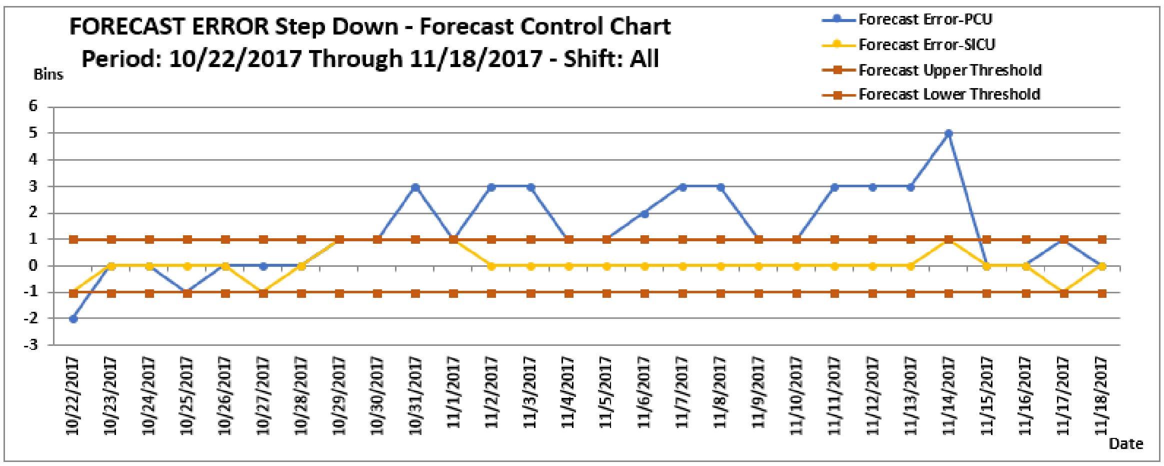
This project has created several artifacts relevant to measuring the individual components of the LPAC cycle. Through in situ interventions, the research team has been able to determine and refine measurements that provide intuitive information to practitioners, allowing for different time series and organizational unit comparisons, and indications for intervention points for corrective or confirming actions. The eADR approach strengthened the results of this project. Each of the final metrics Workload Demand Forecasting (F-*Sb*MAE), Schedule Quality (Completeness, Commitment, Healthiness, Preferences), and Staffing (S-*Sb*MAE) were developed and refined by a combination of researchers and practitioners focused on ensuring that the artifacts met all of the acceptance criteria. The metrics were stress tested with real-world data and then analyzed by the researchers and practitioners to determine usefulness. The practitioners included department leaders from the same departments were the data was drawn providing an additional level of verification and rigor to the analysis.

There are a few limitations to this analysis. There is the possibility that other hospitals may not be collecting the required data for the outcome metrics, but the combined researcher/practitioner team felt that the data chosen would be fairly simple to begin collecting. An additional potential limitation is that this study was conducted entirely within one for-profit health system. It remains to be seen whether these concepts are generalizable other hospitals or non-profit organization hospitals. Non-profit systems may have different approaches to labor-management based on differing missions and organizational focus.

While the metrics used in this study proved to be useful by the practitioner participants, there is no evidence offered concerning the exclusivity of these metrics. Additionally, the project scope and timeline provided no opportunity to determine the long-term sustainability of these metrics as useful measures. Several concepts may apply in these situations that could conceivably weaken the usefulness of the metrics. One such concept is Goodhart's law which indicates that when a goal exists, people will tend to work toward and optimize the one goal regardless of the consequences. A second concept is McNamara's Fallacy which indicates risk in making decisions based solely on easily quantified metrics and ignoring more difficult to quantify observations. Each of these concepts may prove troublesome with the use of the accepted metrics on a long-term basis. Future research needs to determine if the operationalization of these metrics causes employees to develop mechanisms for gaming the numbers to show improvement or the risks involved in focusing on these metrics with respect for overall performance. Future research needs to determine if the prolonged focus on these metrics leads to the intended and desired behaviors improving performance. Each of these concepts will make for interesting future studies.

The next steps in this research are to utilize these metrics in a transformation project to determine if performance outcomes are improvable with a new LPAC management model that focuses on a systems approach to the problem including deep knowledge sharing across silos.





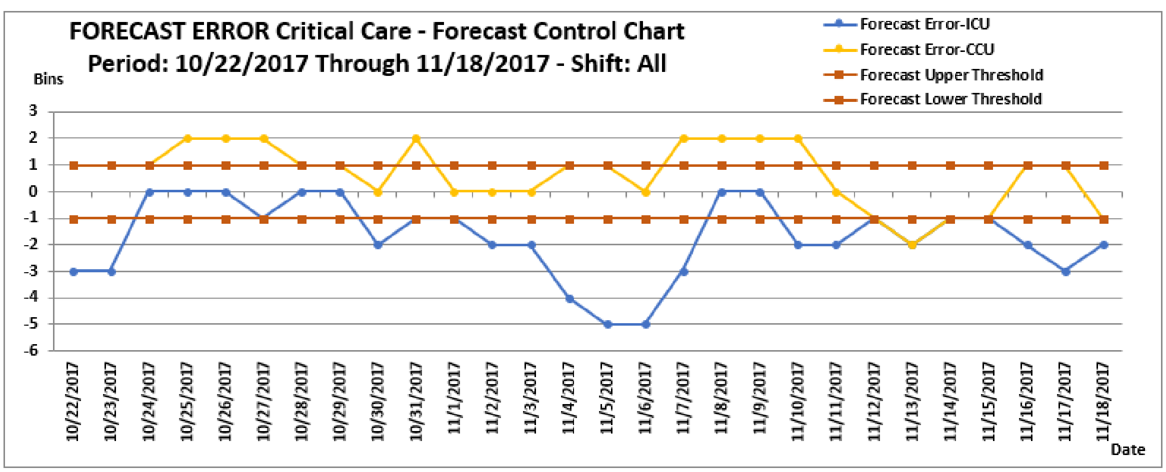
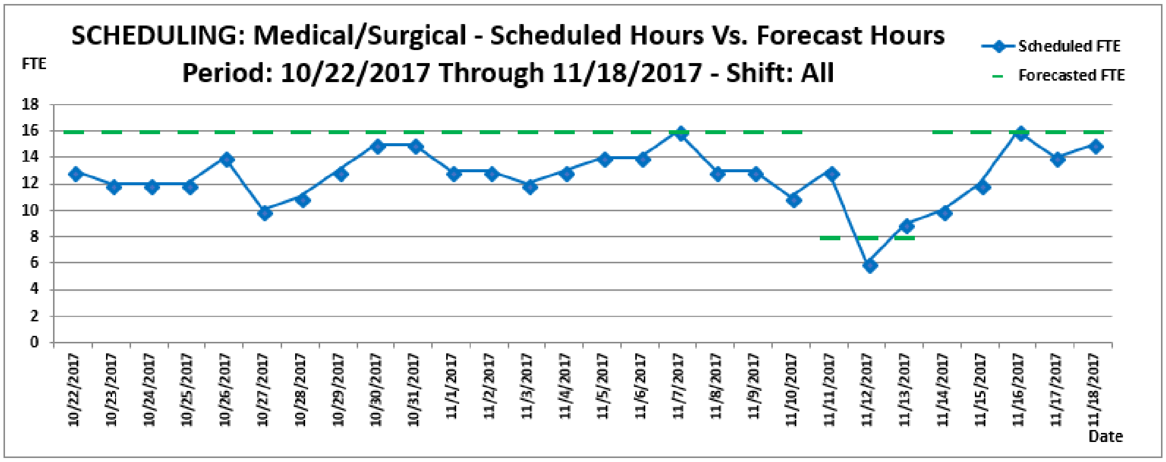
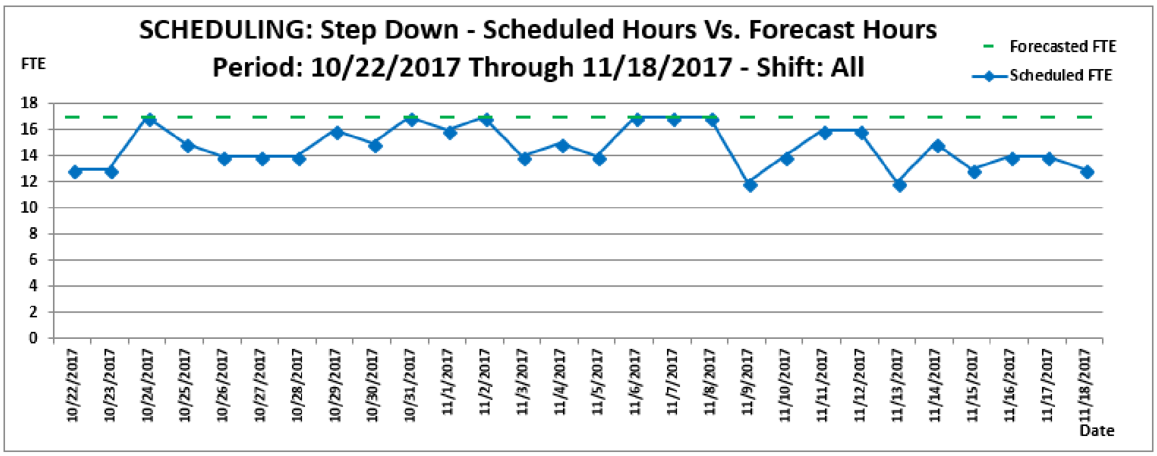


Figure 9 - Service Line Forecasting Control Charts





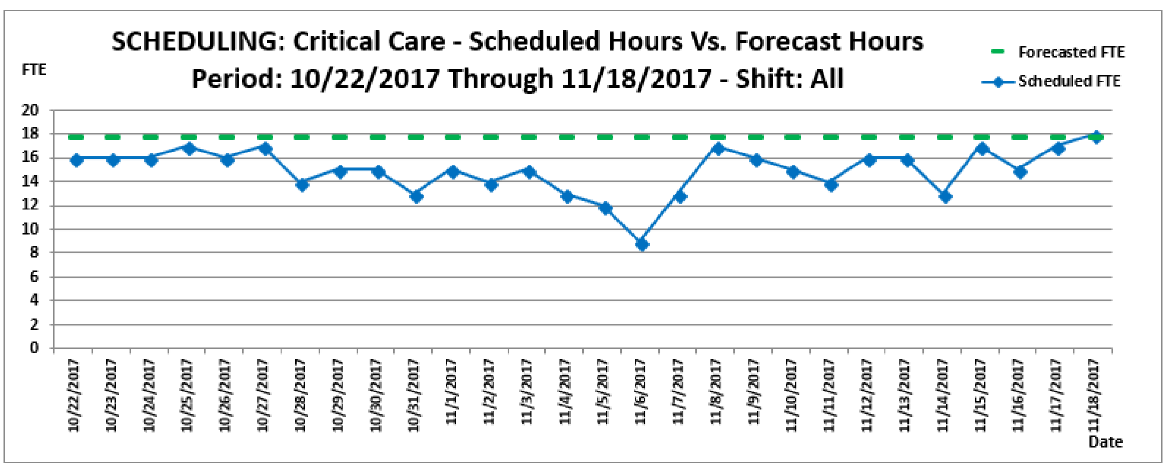
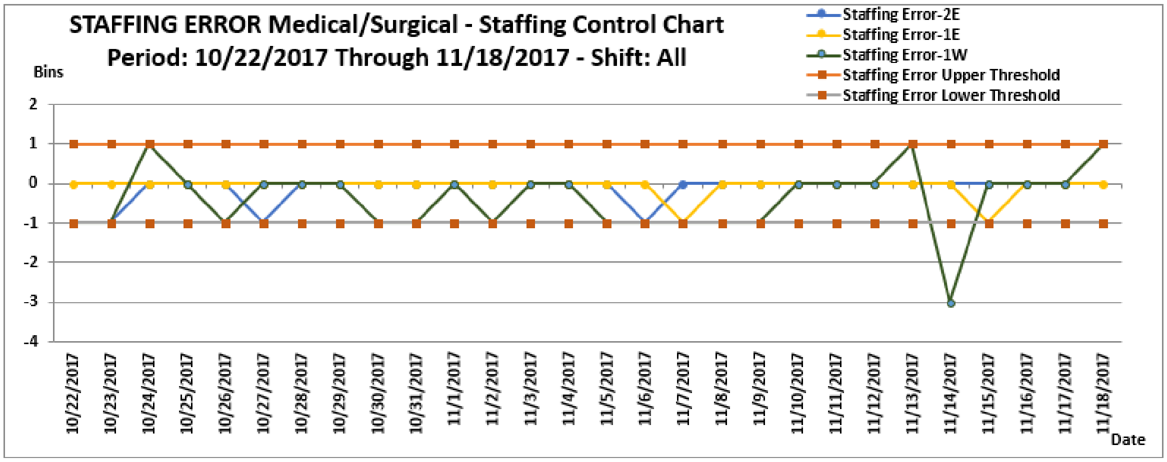
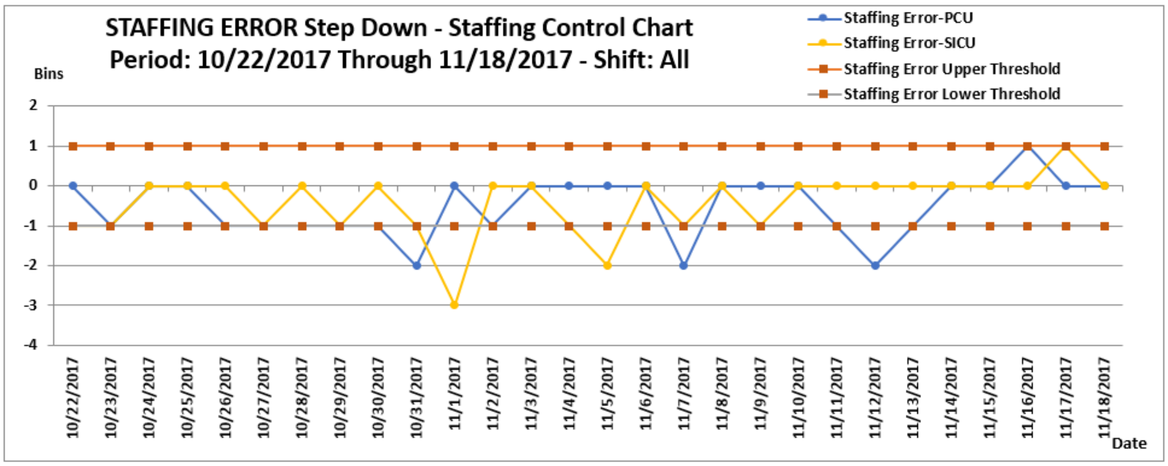


Figure 10 - Service Line Scheduling Completeness Charts





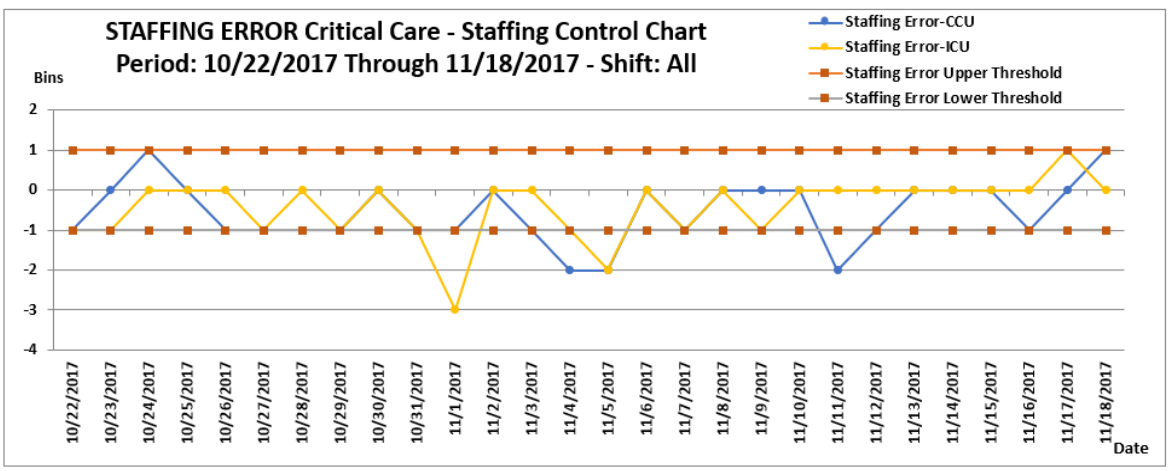


Figure 11 - Service Line Staffing Control Charts

References

Aiken, L. H. (2002). Hospital Nurse Staffing and Patient Mortality, Nurse Burnout, and Job Dissatisfaction. *Jama*, *288*(16), 1987. https://doi.org/10.1001/jama.288.16.1987

Bard, J. F., & Purnomo, H. W. (2005). Short-term nurse scheduling in response to daily fluctuations in supply and demand. *Health Care Management Science*, *8*(4), 315–324. https://doi.org/10.1007/s10729-005-4141-9

Broyles, J. R., Cochran, J. K., & Montgomery, D. C. (2010). A statistical Markov chain approximation of transient hospital inpatient inventory. *European Journal of Operational Research*, *207*(3), 1645–1657. https://doi.org/10.1016/j.ejor.2010.06.021

Cimiotti, J. P., Aiken, L. H., Sloane, D. M., & Wu, E. S. (2012). Nurse staffing, burnout, and health care-associated infection. *American Journal of Infection Control*, *40*(6), 486–490. https://doi.org/10.1016/j.ajic.2012.02.029

Cylus, J. D., & Dickensheets, B. A. (2008). Hospital multifactor productivity: a presentation and analysis of two methodologies. *Health Care Financing Review*, *29*(2), 49–64. https://doi.org/hcfr-29-2-049 [pii]

Defraeye, M., & Van Nieuwenhuyse, I. (2016). Staffing and scheduling under nonstationary demand for service: A literature review. *Omega (United Kingdom)*, *58*, 4–25. https://doi.org/10.1016/j.omega.2015.04.002

Harms, R. (2009). Workforce deployment - A critical organizational competency. *Healthcare Management Forum*, *22*(3), 6–9. https://doi.org/10.1016/S0840-4704(10)60095-6

Holden, R. J., Carayon, P., Gurses, A. P., Hoonakker, P., Hundt, A. S., Ozok, A. A., & Rivera-Rodriguez, A. J. (2013). SEIPS 2.0: a human factors framework for studying and improving the work of healthcare professionals and patients. *Ergonomics*. Taylor & Francis. https://doi.org/10.1080/00140139.2013.838643

Hulshof, P. J. H., Kortbeek, N., Boucherie, R. J., Hans, E. W., & Bakker, P. J. M. (2012). Taxonomic classification of planning decisions in health care : a structured review of the state of the art in OR / MS. *Health Systems*, *1*(2), 129–175. https://doi.org/10.1057/hs.2012.18

Hyndman, R. (2006). Another Look At Forecast-Accuracy Metrics for Intermittent Demand. *Foresight: The International Journal of Applied Forecasting*, (4), 43–46. https://doi.org/10.1016/j.ijforecast.2006.03.001

Johnson, R. L., Tsiros, M., Lancioni, R. A., Johnson, R. L., Tsiros, M., & Lancioni, R. A. (2006). Measuring service quality : a systems approach. https://doi.org/10.1108/08876049510100272

Kane, R. L., Shamliyan, T. A., Mueller, C., Duval, S., & Wilt, T. J. (2007). The Association of Registered Nurse Staffing Levels and Patient Outcomes. *Medical Care*, *45*(12), 1195–1204. https://doi.org/10.1097/MLR.0b013e3181468ca3

Katz, D., & Kahn, R. L. (1978). *The social psychology of organization* (2nd ed.). New York.

Kortbeek, N., Braaksma, A., Burger, C. A. J., Bakker, P. J. M., & Boucherie, R. J. (2015). Flexible nurse staffing based on hourly bed census predictions. *International Journal of Production Economics*, *161*, 167–180. https://doi.org/10.1016/j.ijpe.2014.12.007

Kortbeek, N., Braaksma, A., Smeenk, F. H. F., Bakker, P. J. M., & Boucherie, R. J. (2015). Integral resource capacity planning for inpatient care services based on bed census predictions by hour. *Journal of the Operational Research Society*, *66*(7), 1061–1076. https://doi.org/10.1057/jors.2014.67

Kuhar, P. A., Miller, D., Spear, B. T., Ulreich, S. M., & Mion, L. C. (2004). The Meaningful Retention Strategy Inventory: a targeted approach to implementing retention strategies. *The Journal of Nursing Administration*, *34*(1), 10–18. https://doi.org/10.1097/00005110-200401000-00005

MacPhee, M., Dahinten, V., & Havaei, F. (2017). The Impact of Heavy Perceived Nurse Workloads on Patient and Nurse Outcomes. *Administrative Sciences*, *7*(1), 7. https://doi.org/10.3390/admsci7010007

Maenhout, B., & Vanhoucke, M. (2013a). An integrated nurse staffing and scheduling analysis for longer-term nursing staff allocation problems. *Omega (United Kingdom)*, *41*(2), 485–499. https://doi.org/10.1016/j.omega.2012.01.002

Maenhout, B., & Vanhoucke, M. (2013b). Analyzing the nursing organizational structure and process from a scheduling perspective. *Health Care Management Science*, *16*(3), 177–196. https://doi.org/10.1007/s10729-013-9222-6

Makridakis, S. (1993). Accuracy concerns measures : theoretical and practical concerns. *International Journal of Forecasting*, *9*, 527–529. https://doi.org/http://dx.doi.org/10.1016/0169-2070(93)90079-3

Mark, B. A., Harless, D. W., McCue, M., & Xu, Y. (2004). A longitudinal examination of hospital registered nurse staffing and quality of care. *Health Services Research*, *39*(2), 279–300. https://doi.org/10.1111/j.1475-6773.2004.00228.x

Mullarkey, M. T., & Hevner, A. P. (2015). Entering Action Design Research. In Donnellan B (Ed.), *DESRIST* (pp. 121–134). LNCS 9073.

Needleman, J., Buerhaus, P., Mattke, S., Stewart, M., & Zelevinsky, K. (2002). Nurse-Staffing Levels and the Quality of Care in Hospitals. *New England Journal of Medicine*, *346*(22), 1715–1722. https://doi.org/10.1056/NEJMsa012247

Needleman, J., Buerhaus, P., Pankratz, V. S., Leibson, C. L., Stevens, S. R., & Harris, M. (2011). Nurse Staffing and Inpatient Hospital Mortality. *New England Journal of Medicine*, *364*(11), 1037–1045. https://doi.org/10.1056/NEJMsa1001025

Oldenkamp, J. H. (1996). *Quality in fives : on the analysis, operationalization and application of nursing schedule quality*. University of Groningen. Retrieved from https://www.mendeley.com/research-papers/quality-fives-analysis-operationalization-application-nursing-schedule-quality/?utm_source=desktop&utm_medium=1.17.10&utm_campaign=open_catalog&userDocumentId=%7Be818a5a8-b78d-4b03-aa37-2586200f197b%7D

Park, S. H., Blegen, M. A., Spetz, J., Chapman, S. A., & De Groot, H. A. (2015). Comparison of nurse staffing measurements in staffing-outcomes research. *Medical Care*, *53*(1), e1–e8. https://doi.org/10.1097/MLR.0b013e318277eb50

Silvestro, R., & Silvestro, C. (2000). An evaluation of nurse rostering practices in the National Health Service. *Journal of Advanced Nursing*, *32*(3), 525–535. https://doi.org/10.1046/j.1365-2648.2000.01512.x

Sjögren, K., Fochsen, G., Josephson, M., & Lagerström, M. (2005). Reasons for leaving nursing care and improvements needed for considering a return: A study among Swedish nursing personnel. *International Journal of Nursing Studies*, *42*(7), 751–758. https://doi.org/10.1016/j.ijnurstu.2004.11.001

Tarpey, R. J., & Nelson, M. F. (2009). Schedule Quality Assessment Metrics. *The Health Care Manager*, *28*(2), 145–158.

Thompson, G. M. (1993). Accounting for the multi-period impact of service when determining employee requirements for labor scheduling. *Journal of Operations Management*, *11*(3), 269–287. https://doi.org/10.1016/0272-6963(93)90004-9

Thompson, G. M. (1995). Labor scheduling using NPV estimates of the marginal benefit of additional labor capacity. *Journal of Operations Management*, *13*(1), 67–86. https://doi.org/10.1016/0272-6963(94)00012-4

Thompson, G. M. (1998). Labor Scheduling, Part 1 Forecasting Demand. *The Cornell Hotel and Restaurant Administration Quarterly*, *39*(5), 22–31. https://doi.org/10.1016/S0010-8804(98)89034-4

Wright, P. D., & Mahar, S. (2013). Centralized nurse scheduling to simultaneously improve schedule cost and nurse satisfaction. *Omega (United Kingdom)*, *41*(6), 1042–1052. https://doi.org/10.1016/j.omega.2012.08.004

Yildirim, D., & Aycan, Z. (2008). Nurses' work demands and work-family conflict: A questionnaire survey. *International Journal of Nursing Studies*, *45*(9), 1366–1378. https://doi.org/10.1016/j.ijnurstu.2007.10.010

Chapter Five: Case Study-Human interaction management impact on hospital labor planning

Abstract: This study takes a novel approach to the hospital workforce planning (forecasting and scheduling) problem. The study differentiates itself from precedent work in its focus on the "art" of labor planning (i.e., human interactions and human work within the labor planning processes). Hospital labor planning involves many dimensions and includes a large amount of complexity causing us to believe that many improvement opportunities exist. We did not, however, focus on pure process and information technology tools. Instead, the researchers in this study focused on the human processes, interactions and work involved with forecasting workload and subsequent labor scheduling to redesign necessary components to optimize human interactions, the flow of information, and knowledge sharing to address the large amounts of complexity and variability. The study concluded that a centralized role-process structure that facilitates and encourages more interactions and feedback across the different roles resulted in more accurate labor forecasts subsequently leading to more accurate labor schedules. We found that large amounts of critical knowledge and information became locked within human role participants who did not interact with other roles. There was a lack of a path for the critical information to flow across the roles for successfully performing tasks. The drivers for the improvements were task focus and more information sharing leading to a richer collection of information and knowledge used as input to the work tasks. Redesigning work activities and roles led better forecasting and scheduling outcomes as well as an additional benefit of freeing up clinical department leader time to focus on more patient and employee-centric tasks within their departments.

Keywords: Action design research (ADR), Labor management, Scheduling, Staffing, Human Interaction Management

Note to Reader

Portions of this chapter have been previously published in The Muma Business Review, 2017, 1(11): 125-139, and have been reproduced with permission from The Muma College of Business

Introduction

Over the last three decades, technology has brought many improvements to the efficiency and effectiveness of business processes in almost all industries. Information technology companies developed an intense focus on process facilitation and management with the start of the first enterprise resource management systems in the early 1990's and had continued to innovate in this space for the last twenty-five years. The same has been true concerning hospital workforce planning systems that have likewise evolved toward more complexity attempting to automate processes and solutions to labor planning challenges. While prior research has demonstrated clear links between information technology and process improvement, a large gap has formed between the understanding of how technology facilitates processes and how human interactions affect these same processes. Technology tends to focus on pre-defined, repetitive, sequenced tasks (Lee, Yoon, Seo, Kim, & Kim, 2011). This focus has evolved with the specific objective of automating people out of the process through automating tasks. The same trend has occurred within academic research over the same period. When considering the labor planning process of forecasting and scheduling labor, the large majority of research has centered on the development of new methodologies, mathematical models, or heuristic models to automate the forecasting and scheduling processes. Although many industries have moved to fully automated labor scheduling models facilitated by large, expensive information technology solutions, many healthcare entities have not been able to follow the same path. Despite the overwhelming amount of research toward attempting to optimize the solutions, there have been very few viable solutions operationalized based on this research.

The challenge of labor planning in hospitals is well documented in precedent research typically referred to as the “nurse scheduling” or “nurse staffing” problem. Forecasting labor needs and then matching these forecasted labor needs with actual needs on a shift by shift basis based on the number of patient and severity of illness of patients is a complex undertaking with one common thread, variability. The subsequent inefficiencies and negative impacts resulting from poor scheduling have also been well documented (Cline, Reilly, & Moore, 2003; Holtom & O'Neill, 2004; Vanhoucke & Maenhout, 2009; Wright & Mahar, 2013). The majority of these studies, however, have attacked the problem by proposing new scheduling and forecasting methodologies to be incorporated into either existing or new tools. Organizations already have many technical tools to solve the labor allocation problem, yet they still perform poorly.

This case study considers the challenge of labor forecasting and planning. Workload demand forecasts, which are generally considered the first step of the labor scheduling process, are typically the basis for labor. (Defraeye & Van Nieuwenhuyse, 2016; Ernst, Jiang, Krishnamoorthy, Owens, & Sier, 2004). If the forecasts are accurate, then employees who perform the scheduling function have good information to base their scheduling work and schedule the correct amount of labor. If the forecasts are not accurate, then these same schedules will provide incorrect amounts of labor requiring more effort to be expended in the staffing function to adjust labor to meet the correct levels of demand. Last minute adjustment of labor typically results in higher costs in the form of premium pay (overtime, incentive pay, or contract labor) to fill last minute needs.

On the other hand, last-minute adjustment of labor to send unneeded staff home serves as a job dissatisfier. Staff who expect to work only to be sent home without work or pay leads to irregular work schedules that have been shown to complicate staff work-life balance efforts (Yildirim & Aycan, 2008). Based on the many potential impacts of poor labor schedules, accuracy is a key success factor.

The labor scheduling process within hospitals remains a variable, human-driven process. The process is complex due to a large number of internal and external variables (Siferd & Benton, 1992). These variables can be employee-specific such as differing skill sets, preferences, constraints, individual employee certifications and licensures, experience, and teamwork effects or clinical specific such as patient acuity (severity of illness), patient needs, family member needs, and physician orders. Fully automated solutions that attempt to evolve the process into a pure science of mathematical formulas and calculations have proven elusive to operationalize (De Bruecker, Van den Bergh, Beliën, & Demeulemeester, 2015). Further, De Bruecker et al. conclude that "it is almost impossible to solve problems of realistic size to optimality. This is however not always necessary because the management of a company often prefers a fast and good solution to the optimal solution. Hence, it is not surprising that researchers who are concerned with realistic problems resort to heuristic solution methods instead of exact approaches."

In this case study, we proposed that human-driven processes and interactions including human creativeness concerning how employees should be scheduled and staffed more heavily impact labor allocation processes rather than the tools and software utilized. Hospital labor planning, as indicated above, is a highly variable environment and therefore does not lend well to heavily standardized processes designed to reduce variability. Much like the example given by Hall & Johnson with Ritz Carlton, the effort to document and pre-determine every response and course of action to be taken within the scope of labor forecasting and scheduling in a hospital would be futile (Hall & Johnson, 2009). Instead, we believe successful solutions in this space are dependent on human creativity as generated through human interactions and knowledge sharing. The objective of the study was to analyze the existing labor forecasting and scheduling processes in the hospital within the context of human interaction management and then seek to improve the process by implementing new structures, roles and processes to facilitate critical human interactions and knowledge sharing to foster an environment to support creativity.

In exploring this proposition, an effort was undertaken to embed within one hospital for a 3-month period to first analyze the existing processes and work involved with labor forecasting and scheduling within the lens of the

Human Interaction Management (HIM) framework. HIM includes a focus on human interaction facilitation allowing for knowledge sharing and support of the "art" of labor planning. The framework was used to formulate a new labor-management model to leverage the advantages of human interactions, and finally implement the new model to determine if outcomes could be improved. Before describing the case, the framework a summary of the framework is provided below.

Human Interaction Management

One of the key elements of this study was to move beyond the analysis of business processes and the mathematical models used to forecast labor demand or match open slots in a schedule to specific employees based on multiple constraints. We believed more impactful improvements were possible through an understanding of how the actors within the processes interacted with each other, interacted with the tools and how human activity/creativity impacted the work. We believed there was a great deal of missed opportunity in having the right amount of attention and the right resources focused on specific tasks of labor allocation planning. Additionally, we believed that people performed these tasks less optimally without the facilitation of information sharing across multiple functions.

The Human Interaction Management (HIM) from Harrison-Broninski is an extension of existing Role Activity Theory and Activity Theory. HIM is more of an orienting framework than a theory which typically presents predictive models based on input variables. The concepts are used in this case study by the researchers to anchor discussions and analysis within a complex environment. HIM provides frameworks within the concept that "human activity whether collaborative or not fits into specific patterns and that learning is the core of any collaborative activity" (Harrison-Broninski, 2005). At the core of learning is the sharing of knowledge. While technology exists to store and share knowledge, there is a gap with technology built around automated processes as the sole mechanism for the exchange. HIM proposes that the large deficiency in existing business process management is the neglect for support of human interaction and human learning represented as the "white space" work or the work that occurs in between and around the steps of a formalized process. Documented business processes fall short in their ability to determine how human work gets done within and between each process step. The most thoroughly documented business process fails to fully describe how each actor in the process goes about accomplishing individual tasks (i.e., data gathering, thought processes, analysis, reasoning, organization, etc.). These areas are where we believe the critical elements of labor forecasting and scheduling occur. The very nature of labor-management involves a complex system of actions and reactions which can be difficult to translate into variables and formulas. Each of these actions can be non-conforming to any action previously taken and therefore may not lend well to mathematical algorithms. Instead, the critical path is the understanding of how people interact within the system to readily adapt to real-time changes. Much of the information required exists not stored in computers but rather within human participants who can make sense of the rapidly changing variables. We proposed that forecasting and scheduling include both "science" and "art" for successfully providing accurate workforce planning.

HIM includes five principles: (1) team building – creating effective teams including role definitions, (2) communication – structured and goal-directed in order to manage interactions, (3) knowledge – learning to manage time and mental effort, (4) empowered time management – understanding how humans structure work, (5) collaborative, real-time planning – understanding how humans make things happen (Harrison-Broninski, 2006). We considered each of these principles in our structure of the involved tasks to document critical interactions and human work.

Harrison-Broninski also proposed a five-stage model for how humans work (REACT) (Harrison-Broninski, 2005):

- Research – information and knowledge gathering
- Evaluate – consider the knowledge gained
- Analyze – decide on an approach based on a new understanding
- Constrain – divide work into manageable chunks
- Task – accomplish the tasks

This model was critical to our understanding of the structure of how human work was approached and completed by the participants to seek areas where improvement opportunities existed. The work and interaction involved within the processes were documented utilizing the Role Activity Diagram (RAD) methodology (Ould, 1995; Phalp, Hendersonb, John, & Abeysinghe, 1998) and refined by Harrison-Broninski (Harrison-Broninski, 2005). RAD is an intuitive tool used to model business processes first dating back to the 1980's that is understandable to the process participators. The tool has been refined over the last three decades to provide a tool that intended to "show how people work together to accomplish their individual and shared goals" (Harrison-Broninski, 2005) and therefore was an ideal framework for anchoring our discussions and research. We hoped that we could leverage this framework to analyze existing work structures and processes to pinpoint weaknesses where roadblocks to information sharing and human interaction existed. We proposed that these weaknesses could be overcome through the restructuring of the processes and work to remove the roadblocks and facilitate the human interactions and knowledge sharing to result in improved outcomes in the form of more accurate workload forecasts and higher quality labor schedules.

The Case

This study was undertaken at a medium-sized hospital in the Midwestern United States and lasted for four months involving three departments. After previously seeking labor cost improvements through implementing labor schedule quality reviews, the facility chose to continue the effort to investigate controls on labor costs. The leadership of the hospital had identified rising labor costs and negative feedback from staff concerning schedules and employee satisfaction. The main thread of the feedback was centered on employees rarely working the actual days they were scheduled making it difficult to plan personal lives. The facility sought to address these issues without a large technology replacement that was not practical from a cost or a resource perspective after not finding much evidence of significant improvements from other facilities who had implemented such systems.

The study consisted of the following steps: pre-implementation assessment, design, implementation, testing, and post-implementation assessment. The pre-implementation assessment consisted of direct observation, participation, and interviews with the Chief Financial Officer, department directors and department managers (decentralized model) during the creation of the labor forecast and 2-week schedules for multiple departments. From data gathered and analyzed, role activity diagrams were redesigned and then leveraged for improvement opportunities. The role activities were improved to facilitate interactions and knowledge sharing and subsequently documented. The next step was the implementation of the new model for the creation of another labor forecast and a 2-week schedule.

Measurement Criteria

The process of labor scheduling can be broken down into two critical work packages. A quantified workload demand (forecast) is necessary to understand how many staff members to schedule. A scheduling resource then creates a schedule with the names of the employees assigned on a per shift basis based on the pre-determined staffing requirements required to meet the forecasted workload (Defraeye & Van Nieuwenhuyse, 2016).

The measurement of these two variables was completed via schedule quality and labor forecast accuracy metrics as calculated using data from existing systems.

The concept of schedule quality included three components used to measure the quality of a labor schedule as detailed in Table 19 (Tarpey & Nelson, 2009). These metrics were used to compare the quality of the schedules created pre-implementation to the schedules created post-implementation to measure potential performance improvements. IT tools were available for scheduling providing the data for the assessment of schedule quality based on three dimensions: completeness viewed from the perspective of all staff, nurses (professional), and non-nurses (support), commitment, and healthiness. These dimensions were used to assess the labor schedule's ability to map to expected demand, fully pre-allocate all labor, and provide healthy work patterns absent of long hours or long stretches of continuous days scheduled. A schedule was essentially complete when the amount of labor required to meet a forecast workload demand was scheduled. A schedule had a full commitment when all resources on a schedule were scheduled to the full allotment of hours available to work. A schedule achieved healthiness when there was an absence of generally accepted unhealthy work patterns in the schedule such as consecutive 12-hour work shifts, short rest periods between shifts, etc. (Tarpey & Nelson, 2009).

Table 19 – Schedule Quality Components

Component	Definition	Metric
Schedule to Department Needs	Measures the schedule's prospective effectiveness and efficiency toward meeting department demand (provision of staff for a projected volume)	Total Completeness (Effectiveness)
		Professional Completeness (Effectiveness)
		Support Completeness (Effectiveness)
		Commitments (Efficiency)
		Healthiness (Satisfaction)

The forecast accuracy measurement SbMAE (Staffing Bin Mean Average Error) measures the workload forecast for accuracy. Nursing departments at this facility based their scheduling on staffing matrices. A staffing matrix is a staffing plan that indicates how many of each type of employee (skill) is required to care for a pre-determined range of patients. The staffing matrices were determined annually during the budgeting process. Hospital leadership worked with departmental leadership to determine how many of each skill set to hire in the department and how those resources should be allocated based on patient load all falling within the parameters of the overall annual budget. The result was a staffing plan (Staffing Grid) stored in Excel documents on the hospital Intranet. The process was extremely important as the number of each skill set allowed for each range of patients had a direct impact on the quality of patient care. The underestimation of the number of RN nurses needed would most likely cause patients not to receive the care they need as nurses would have too large of a workload. An overestimation of resources resulted in more labor in the unit then was needed causing overspending on labor. Table 20 provides an example of one of the staffing grids.

Table 20 – Critical Care Department Staffing Grid Example

Patient Minimum	Patient Maximum	Charge-RN	RN	Nurse Tech	Unit Clerk	Labor Bin
0	**0**	0	0	0	0	1
1	**2**	1	1	0	0	2
3	**4**	1	2	0	0	3
5	**6**	1	3	0	0	4
7	**8**	1	4	0	1	5
9	**10**	1	5	0	1	6
11	**12**	1	6	0	1	7
13	**14**	1	7	1	1	8
15	**16**	1	8	1	1	9
17	**18**	1	9	1	1	10

Each row in the example corresponds to a labor bin. A forecast was accurate when the actual number of patients corresponded to a labor bin that is ±1 labor bin from the labor bin that corresponds to the forecasted number of patients. For example, if the forecast number of patents is 9, the corresponding labor bin is #6. An accurate forecast is where the actual number of patients is between 7 and 12. Seven patients represent the bottom of labor bin #5 which is one bin below labor bin #6, and Twelve patients represent the top of labor bin #7 which is one bin above labor bin #6. In each of these cases, the labor change to go 1 labor bin up or down is only 1 employee, which is considered a reasonable last-minute staffing change. SbMAE measures the average error for a given time series to provide an overall accuracy measurement. A SbMAE value between 0 and 1.0, inclusive is considered a forecast within an acceptable tolerance.

Current State Background

Typical of many hospitals in the United States, scheduling at this facility was decentralized down to the unit or department level. Individual department leadership was responsible for the scheduling of labor. This decentralized model of work was analyzed resulting in the role activity diagram detailed in Figure 12. There were three main roles within this process: forecaster, scheduler, and staff. We observed the steps of the process as follows:

1. CFO generated the workload demand forecast
2. CFO input the forecast (made available) into the process utilized by the scheduler
3. Department Directors opened labor schedules for employees to self-schedule their desired shifts
4. Department Directors closed labor schedules for employee input and then modified the schedule accordingly
5. Department Directors either reviewed schedule quality and modified the schedule accordingly or proceeded directly to making the schedule available to employees
6. Department Directors received employee feedback if any
7. Department Directors modified labor schedule based on employee input, if necessary
8. Department Directors finalized and communicated the labor schedule to employees

In this model, the actors who played the various roles were the Chief Financial Officer (CFO) serving in the forecaster role, the Department Directors serving in the scheduler roles, and individual staff members serving in the staff role. The labor demand forecast was generated on a monthly basis by the CFO and pushed down to each department to provide a baseline for the number of each staff type or skill to be scheduled. The basis for the forecast was the annual budget with some consideration for the current run rate from an admissions perspective. The overall hospital admission budget was then broken down to the department level. This approach was a macro to micro approach with an aggregated number broken down into a per department forecast. Each department director was responsible for creating labor schedules that mapped to the department workload forecast, fully allocated all staff to work shifts with a minimum of unhealthy shift patterns and make a best effort to accommodate staff preferences. Several critical findings came out of this observation and mapping effort. There were two areas in the diagram where an expectation existed regarding the occurrence of a robust REACT process of work. In the forecasting work package, we found that the role participant (CFO) hit the five stages, but with a much lower intensity than we expected in the research phase. We gathered observations from the process and created Table 21 to summarize.

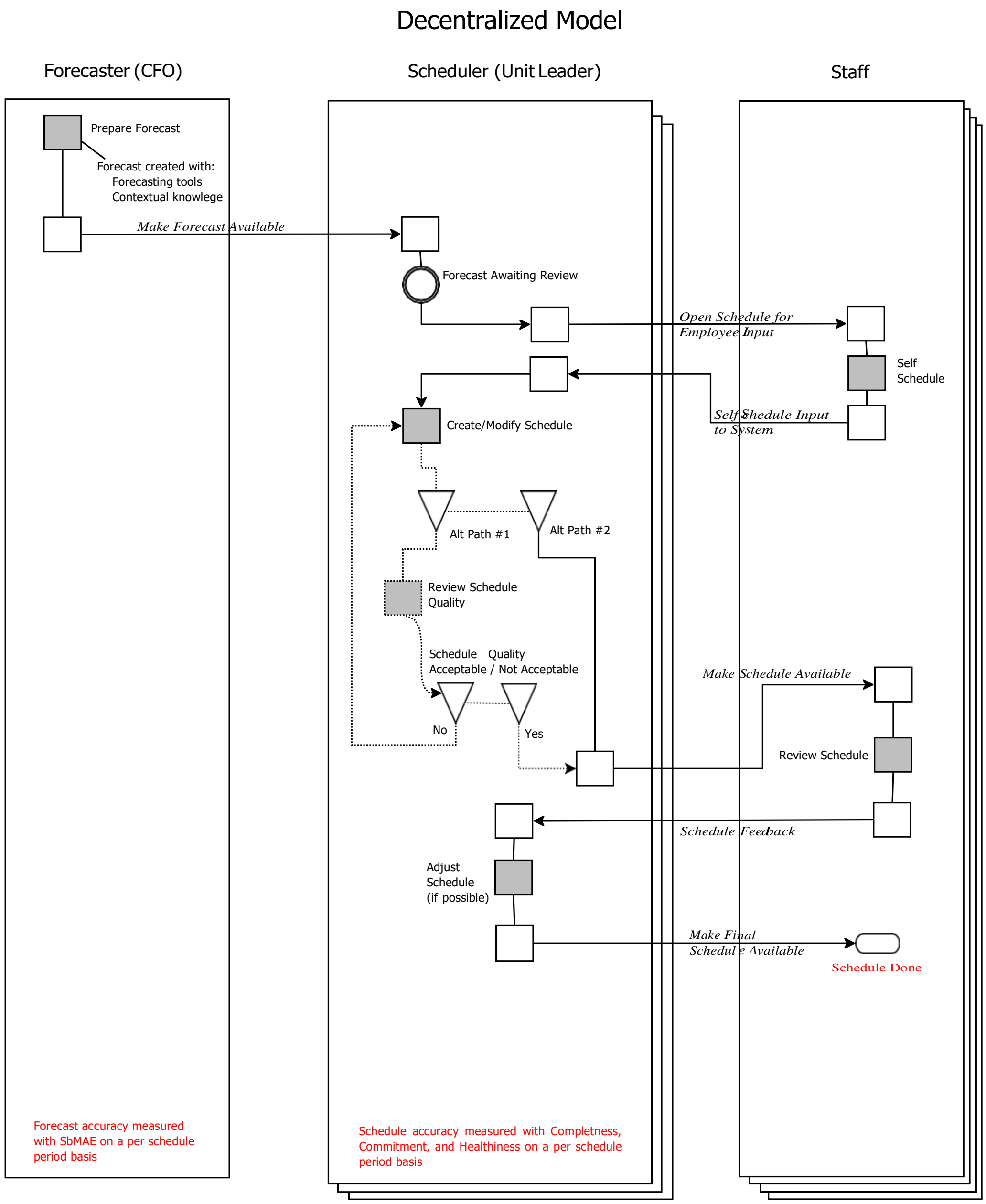


Figure 12 - RAD (Decentralized Model)

The CFO's research was limited to gathering financial information concerning the current trend of admissions compared to the budgeted trend of admissions. There was no effort to gather information regarding the real-time environmental criteria of the department (i.e., what was currently occurring). We expected a more robust effort to understand the current state of each department concerning any existing constraints or volume determinants (i.e., blocked beds, construction, physician/surgeon schedules and vacations, construction, etc.). We noted this item as a critical shortcoming concerning interactions and information sharing. We concluded that too little input data was used to generate forecasts.

We found validation of our concerns in the Decentralized Model RAD (Figure 13) in several areas. The lack of any feedback information flow from the scheduler role to the forecaster role indicated the lack of information sharing concerning the forecast. The two separate paths within the scheduler role (Alt Path #1 and Alt Path #2) indicated the varied approach as to whether or not the scheduler made use of the schedule quality metrics in guiding their scheduling work. Lastly, schedulers relied on the self-scheduling process as the determinate for staff preferences. There was no guarantee that the staff would get the shifts for which they signed up. Once the employee input phase ended, schedulers balanced the schedule (moved employees around) with little additional interaction with staff. Table 22 summarizes observations from the scheduling process.

Table 21 – REACT analysis of Actor in Forecaster Role (Pre-Implementation)

REACT Phase	Expectation	Reality
Research	Investigate the principles, talk to those in the know, locate potential threats to gain information from external sources and turn it into personal knowledge	Limited to an individual gathering of financial information concerning the current trend of admissions compared to the budgeted trend of admissions - limited external contact
Evaluate	Step back and consider the knowledge acquired	Information internalized via repetitive process each month
Analyze	Based on the new found understanding decide how to approach the problem	Due to the repetitive nature of work, approach to the problem is nearly the same each cycle
Constrain	Divide work into separate chunks and define constraints that govern the work	Very little time spent in phase as the work typically defined as one chunk with the only constraint of completing by a certain date
Task	Complete work	Work completed

Table 22 – REACT analysis of Actors in Scheduler Role (Pre-Implementation)

REACT Phase	Expectation	Reality
Research	Investigate the principles, talk to those in the know, locate potential threats to gain information from external sources and turn it into personal knowledge	Limited to acceptance of the forecast with no interaction even if the forecast was deemed inaccurate. Some interaction with staff on preferences outside of self-scheduling
Evaluate	Step back and consider the knowledge acquired	Little time spent considering forecast. Preferences considered mostly on pre-existing knowledge and memory
Analyze	Based on the new found understanding decide how to approach the problem	Due to the repetitive nature of work, approach to the problem is nearly the same each cycle
Constrain	Divide work into separate chucks and define constraints that govern the work	Some directors chunked work by shift delegating to a day and night charge, or the director completed the entire schedule
Task	Complete work	Work completed

Centralized Model

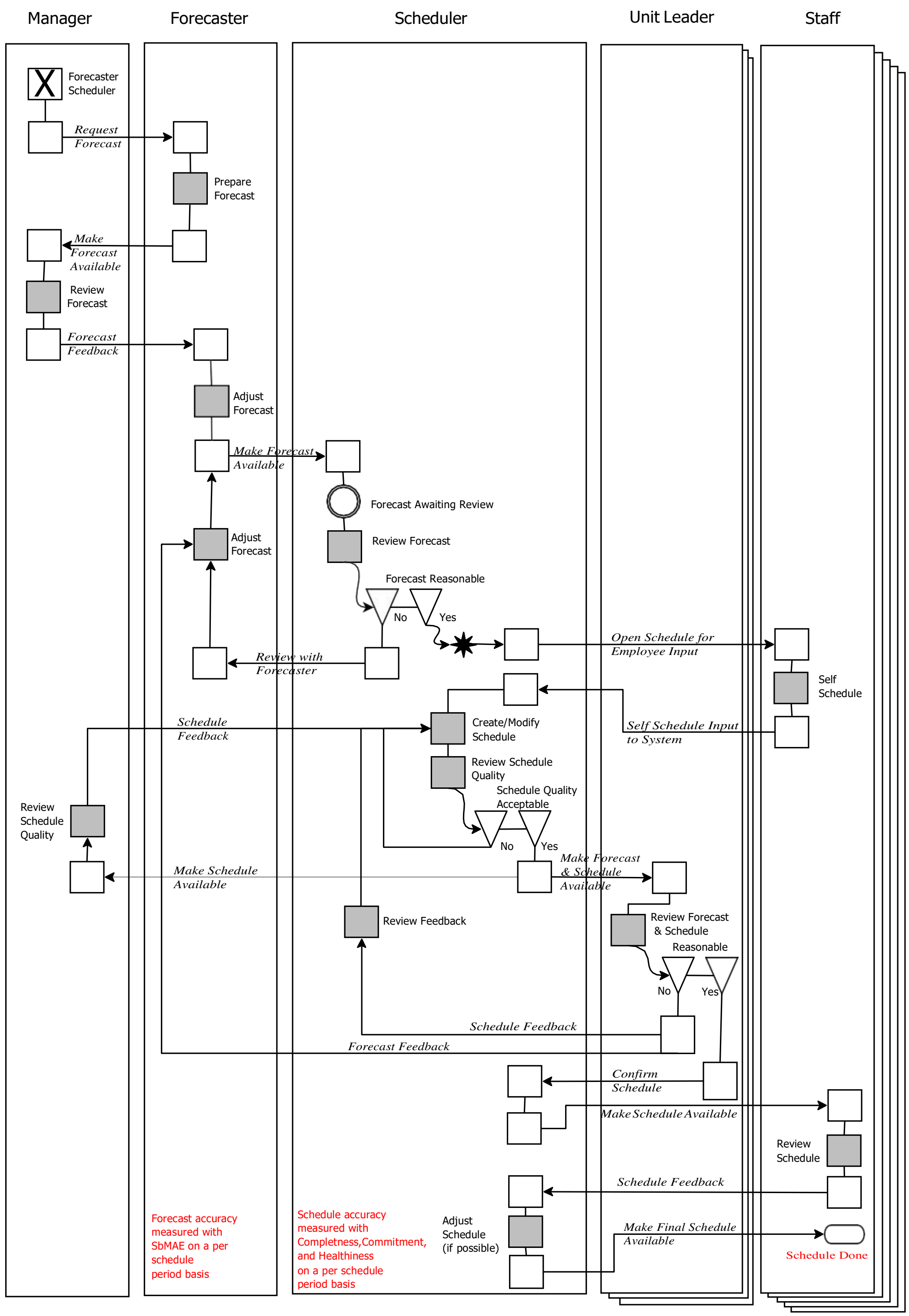


Figure 13 - RAD (Centralized Model)

We summarized our current state review work with two main conclusions fed into the design stage. First, the role participants did not consistently make efforts to gather information and data to internalize into knowledge for creating either the forecast or the schedules. Second, there was very little role interaction in a process that should benefit from heavy interaction and information to execute successfully. We determined that part of the reason for lack of focus in these areas was the quantity and type of work involved. Clinical department leaders and the CFO faced large amounts of critical work during a given work day. The forecasting and scheduling work in question was viewed as less important, mundane tasks that competed for attention directly with more critical tasks.

New Model Design

The design of the new model began with a focus on role interactions. While the decentralized roles remained intact, we looked at which participants served in the roles to determine if the most effective and knowledgeable resources were participating at the optimal points in the process. We also looked at adding additional roles to provide more specialization in task performance. It concerned us to find that critical functions of labor planning obtained only secondary importance status. Labor workload forecasting and scheduling involve a great deal of interaction and information sharing. Hospital environments are real-time oriented where multiple variables can impact forecasts and schedules in short to immediate time frames. Therefore, typically, a combination of system generated and judgmental forecasting techniques are employed. Initially, systems create baseline forecasts, and then human intuition and last-minute knowledge are applied to adjust the forecast. For example, a prominent heart surgeon that decides to take a two-week vacation will have an impact on the patient volume of a heart unit since the number of surgeries during that two weeks will decline. This type of information is not typically found in systems or accounted for in historical extrapolations, but rather is stored solely as human knowledge requiring communication and interaction to be useful. For the new model to provide value, we needed to design a work/interaction process to leverage the full value of this information.

Our approach was to centralize the scheduling activities and move the forecasting activities to a different resource to provide a more focused approach. Rather than relying on secondary attention for task completion, we felt that moving the tasks to primary attention and priority would provide better overall performance. We studied the original RAD and focused on new interaction points to facilitate knowledge sharing and equally important, additional feedback loops. The steps in the new model were:

1. New forecasting position generated the workload demand forecast
2. The forecast reviewed by the manager of the new labor planning department and provides feedback
3. Forecaster adjusts forecast based on any feedback from the manager
4. Forecaster makes the forecast available to the scheduler
5. Scheduler reviews forecast providing feedback to forecaster
6. Scheduler opens labor schedule for employees to self-schedule their desired shifts
7. Forecaster adjusts forecast, if applicable and sends forecast back to the scheduler
8. Scheduler creates labor schedule
9. Scheduler reviews schedule quality modifying the schedule to improve quality
10. The scheduler provides schedule and forecast to the manager and department leader for feedback
11. Forecaster adjusts forecast based on any feedback delivered from department leader
12. Scheduler adjusts the schedule based on any feedback delivered from the manager, department leader or any forecast updates
13. Upon confirmation of labor schedule between the manager, scheduler, forecaster and department leader, the schedule is finalized and communicated to employees

The forecasting and scheduling tasks in this model were both centralized with dedicated resources. The human interactions between the department leadership and the scheduling and forecasting resources, where a large portion of the contextual knowledge existed to feed into both forecasts and schedules, were designed to be enabled at the right times. A comparison of the centralized versus the decentralized models shows that the centralized model includes seven work review points across the five roles where the decentralized model only had two work review points across three roles. The new centralized model expanded the number of roles in the process and presented more opportunities for forecast and schedule feedback and refinement.

The next step was to hold work session meetings with department leaders and department charge nurses to thoroughly review and explain the new roles, responsibilities, and processes. Department leadership had ample opportunity to critique and provide input to the new centralized model. This effort served two purposes. First, the inclusion of department leadership in the process design work led to many productive suggestions to improve the final model. Second, buy-in was generated from these same department leaders as they became part of the work team and subsequently assumed shared responsibility and ownership for the success of the model. The session produced an evaluation plan for the implementation phase including key success metrics such as forecast accuracy measurement: SbMAE and schedule accuracy measurements: completeness, commitment, and healthiness.

The final step of design and build was a work session that included the facility leadership team (including the CEO, CNO, COO, and CFO) to review the new design. We explained the reasoning behind the work assignment structure and discussed at great length the expected improvements.

Forecasting and Scheduling Process Changes

Candidates to fill the various roles were selected, trained, and placed into the new roles. Several weeks of parallel runs were executed to ensure process and tool understanding as well as the opportunity to build relationships between process participants. The processes were then turned over to the new resources serving as manager, forecaster, and scheduler. For the following schedule period, the central forecaster created the labor workload forecast for the three test departments. Figure 14 details the tasks executed with volume forecasting feedback occurring across the three touchpoints indicated by A, B, and C and schedule feedback occurring across the four touchpoints indicated by D, E, F, and G. The most recent, relevant information was shared across multiple roles and taken into account within the forecast. Forecasts subsequently included historical trends as well as judgmental real-time knowledge adjusting the expected patient volumes. Expanding the number of roles and participants provided the ability for a richer collection of knowledge and information for consideration.

The scheduling work package involved heavy interaction between the manager, scheduler and the department leaders concerning review of the schedule for reasonability. While the schedule quality metrics guided the scheduler toward creating accurate labor schedules, increased knowledge sharing between the department leader and the scheduling resource led to the creation of schedules better serving the staff from a satisfaction perspective. The schedules also better served the department from a resource coverage perspective. The scheduler also developed personal relationships with many of the staff resulting in an increased knowledge of preferences via the creation of preference documentation. This documentation provided a more reliable data source regarding staff preferences rather than relying on department leader memory and intuition common in the prior decentralized model.

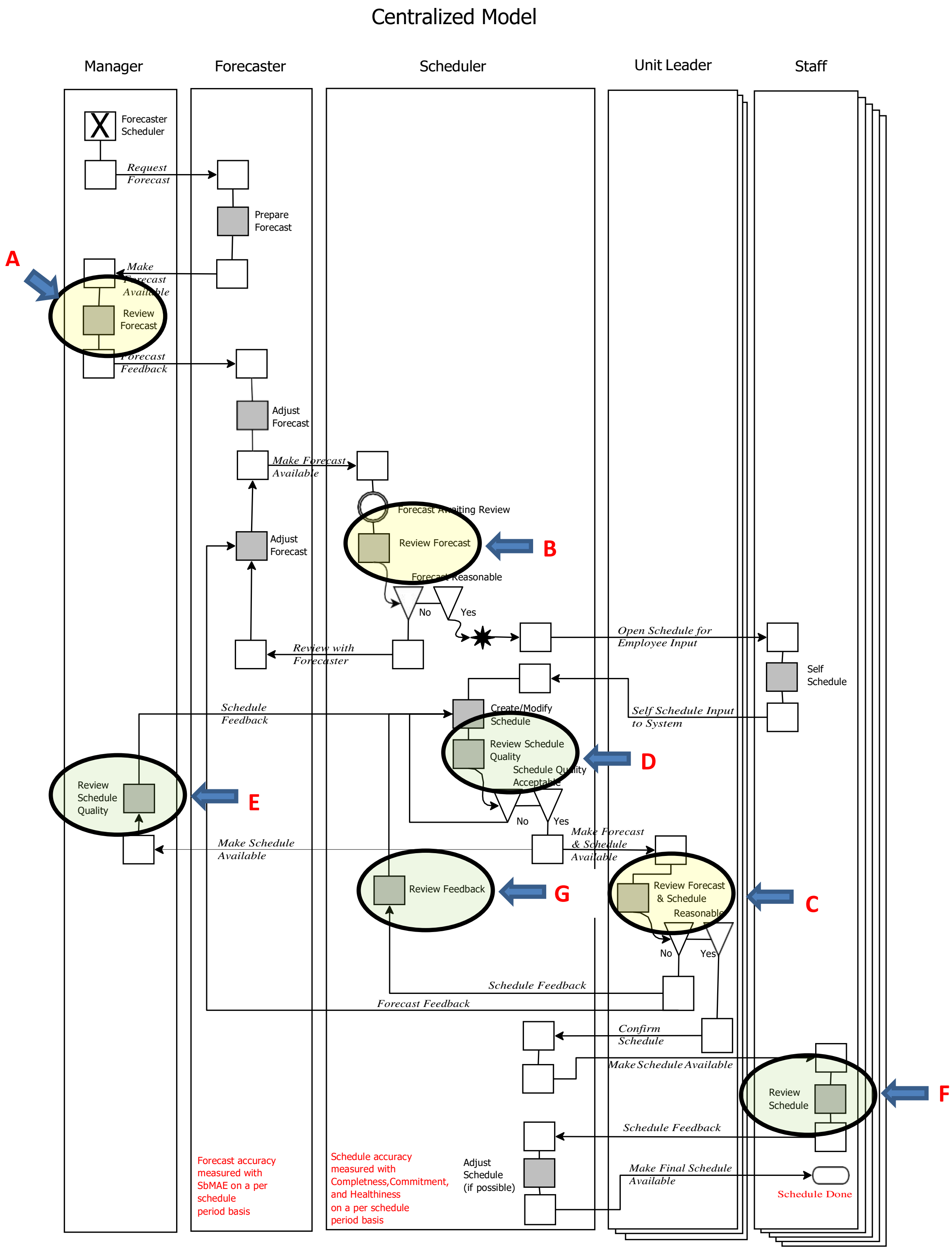


Figure 14 - RAD (Centralized Model) –Review Touchpoints

Discussion

Initial analysis of the centralized model results proved positive. Our original data collection bore out improvement opportunities in forecasting accuracy and schedule quality for the three test departments pre-implementation. These metrics exhibited improvement during our test period post-implementation as detailed in Table 23 and Table 24.

The Staffing Bin Mean Average Error (SbMAE) measurement has an accuracy threshold of 0.00 – 1.00 indicating that the forecast workload was, on the average, within one labor breakpoint or labor bin on the staffing grid resulting in a minimal amount of labor change. A score greater than 1.0 indicates that the labor forecast was sufficiently inaccurate to result in the likelihood of a detrimental last-minute labor adjustment resulting in added costs. Table 23 details SbMAE metric scores pre and post.

Table 23 – SbMAE Scores for Forecasting Before Implementation

	SbMAE Period #1 Pre-Implementation	SbMAE Period #2 Pre-Implementation	SbMAE Period #1 Post-Implementation
Unit A	1.43	2.07	0.41
Unit B	1.21	1.71	0.44
Unit C	1.29	1.39	0.56

The pre-implementation SbMAE scores greater than 1.00 indicate that the forecast produced within the decentralized model was less accurate on the average. The forecast was more than one labor break point off the actual need realized in the department. The forecasting SbMAE scores with the centralized, post-implementation model indicate an improvement in forecasting workload with scores less than 1.00 indicating that the forecasted labor need was closer to the actual labor needed.

The forecasting technique used in the decentralized model was a combination of quantitative and judgmental forecasting. Contextual knowledge of the facility and departments formed the basis for the judgmental forecasting as the participant in the role of forecaster and the manager of the new centralized department were both familiar with the hospital having both worked in the facility for some years. The contextual knowledge based on experience added to the statistical methods and provided a more accurate forecast, consistent with findings from Sanders and Ritzman who concluded that judgmental forecasts based on contextual knowledge combined with statistical forecasts improve forecast accuracy (Sanders & Ritzman, 1995).

Likewise, the schedule quality measurements showed the potential for improvement from the pre-implementation state. Scores less than 90% are considered opportunities for improvement. Table 24 presents the scores for the two pre-implementation periods and the post-implementation periods.

Table 24 – Schedule Quality Scores Before Implementation

	Pre-Implementation						Post-Implementation		
	Completeness Period #1	Completeness Period #2	Commitment Period #1	Commitment Period #2	Healthiness Period #1	Healthiness Period #2	Completeness Period #1	Commitment Period #1	Healthiness Period #1
Unit A	83.16%	81.84%	100%	100%	92.73%	93.40%	94.25%	100%	92.66%
Unit B	78.21%	77.92%	100%	100%	99.01%	99.01%	85.02%	100%	98.51%
Unit C	77.73%	67.86%	100%	100%	90.80%	93.08%	91.55%	100%	94.91%

The completeness metric improved in each of the departments indicating a better mapping to the labor forecast. We propose that the reason for this improvement was the increased interaction between the forecaster, the scheduler, and the department leader. The scheduling resource appeared to have more confidence in the labor forecast and created schedules that mapped closer to the expected need. Since all participants were working from the same assumptions and knowledge, it is logical to conclude that the increased interaction was an improvement. The commitment metric showed no improvement over the pre-implementation state, but this is one area where a focus was already present, and schedules performed well. The implementation of the new model retained the same level of performance in ensuring that all staff was scheduled to their full allocation. Therefore, there was no negative impact with the new model on the commitment metric performance. Lastly, the healthiness metric improved over the first pre-implementation period but was slightly worse from the second pre-implementation period. This metric is directly dependent on the expected patient volumes. If patient volumes are expected to increase, we can expect unhealthier schedules as employees volunteer to cover extra shifts and work overtime. In each of the departments, the volume forecast was higher in the post-implementation period resulting in a higher need for scheduled labor. Most hospitals staff their departments to a certain level with the expectation that additional labor needs will be covered with overtime and contract labor. This staffing methodology allows hospitals to smooth their hiring needs between busy and slow seasons. The post-implementation period used in this study was deeper within this hospital's busier season, and therefore patient volumes were higher. In discussions with the facility leadership, we concluded that the decrease in schedule healthiness was not significant enough to cause concern given the higher patient volumes forecasted.

An additional benefit documented as part of this project was that clinical department leaders indicated they experienced a time savings in being relieved of the labor scheduling task. Interested in this aspect, we surveyed the three leaders and asked them to estimate the amount of time they dedicated to staffing and scheduling before the implementation of the centralized model. The findings resulted in an average of 10-15 hours of time saved on a weekly basis. Additional time was now available for department leaders to focus on more critical tasks involving employees or patients, which previously competed for attention with the labor planning tasks.

Conclusion and Next Steps

We concluded from this study that there is a case to be made for structuring the work of labor forecasting and scheduling in a hospital within a centralized department with dedicated resources for forecasting and scheduling. The success of this structure, however, is directly dependent on the structure of the human work and interactions. Role participants need to have the necessary interactions with other role participants to bring the full value of human stored information and knowledge into the work activities to result in the most informed forecasts and schedules thus allowing for the "art" of the process to be facilitated. The free-flow of information between the roles of manager, forecaster, scheduler, and department leader serve to increase both input and review of work products leading to better quality. Additionally, providing resources that are focused on specific tasks allows for more specialization in skills moving labor forecasting and scheduling functions from a secondary to a primary function allowing for the higher concentration of focus and subsequently better outcomes. The increased focus allows participants to engage more fully within the steps of the REACT process allowing for higher quality work. HIM is a useful framework for analyzing the human activity involved in labor forecasting and scheduling. The concepts provide context for structuring work processes to facilitate human interaction and robust information sharing specifically within highly human-driven tasks that deal with high variability.

The study was limited in the scope of departments and number of schedule periods that were studied. While we used three departments in one facility, more departments are needed to fully assess the impact of the new work and organizational structures. Additionally, research needs to continue across longer time periods as well as within multiple hospitals to verify the positive impact.

There are several next steps for this research. The first important question to answer is whether or not the proposed centralized structure is scalable and if so, what quantity of improvement comes with scale? The second question to answer is whether or not the improved forecasts and labor schedules can be leveraged into better staffing performance to have an impact on labor costs further down the process chain. Lastly, it is important to study whether or not the improvements gained by implementing the new model are sustainable over the long run. Many times, improvements are immediately achieved only to be watered down or lost as time goes on and people revert to their old way of doing things. Each of these questions will serve as valuable research questions for ongoing research into labor planning management models within hospitals.

References

Cline, D., Reilly, C., & Moore, J. F. (2003). What's behind nurse turnover? *Nursing Management, 34*(10), 5.

De Bruecker, P., Van den Bergh, J., Beliën, J., & Demeulemeester, E. (2015). Workforce planning incorporating skills: State of the art. *European Journal of Operational Research, 243*(1), 1-16. doi:10.1016/j.ejor.2014.10.038

Defraeye, M., & Van Nieuwenhuyse, I. (2016). Staffing and scheduling under nonstationary demand for service: A literature review. *Omega, 58*, 4-25. doi:10.1016/j.omega.2015.04.002

Ernst, A. T., Jiang, H., Krishnamoorthy, M., Owens, B., & Sier, D. (2004). An annotated bibliography of personnel scheduling and rostering. *Annals of Operations Research, 127*(1-4), 124.

Hall, J. M., & Johnson, M. E. (2009). When should a process be art, not science? *Harvard Business Review, 87*(3), 9.

Harrison-Broninski, K. (2005). *Human interactions: The heart and soul of business process management*. Tampa, FL: Meghan-Kiffer Press.

Harrison-Broninski, K. (2006). Human interaction management (him) – quick reference card. http://harrison-broninski.com/keith/him/HIM-quick-reference-card.html.

Holtom, B., & O'Neill, B. (2004). Job embeddedness: A theoretical foundation for developing a comprehensive nurse retention plan. *Journal of Nursing Administration, 34*(5), 12.

Lee, J., Yoon, J., Seo, W., Kim, K., & Kim, C.-H. (2011). A fact-oriented ontological approach to human process modeling for knowledge-intensive business services. *Expert Systems with Applications, 38*(10), 12281-12292. doi:10.1016/j.eswa.2011.04.006

Ould, M. A. (1995). *Business processes: Modelling and analysis for re-engineering and improvement*. California: Wiley Inc.

Phalp, K. T., Hendersonb, P., John, W. R., & Abeysinghe, G. (1998). RolEnact role-based enactable models of business processes. *Information and Software Technology, 40*(1998), 11.

Sanders, N. R., & Ritzman, L. P. (1995). Bringing judgment into combination forecasts. *Journal of Operations Management, 13*(4), 11.

Siferd, S. P., & Benton, W. C. (1992). Workforce staffing and scheduling: Hospital nursing specific models. *European Journal of Operational Research, 60*, 233-246.

Tarpey, R. J., & Nelson, M. (2009). Schedule quality metrics. *The Health Care Manager, 28*(2), 13.

Vanhoucke, M., & Maenhout, B. (2009). On the characterization and generation of nurse scheduling problem instances. *European Journal of Operational Research, 196*(2), 457-467. doi:10.1016/j.ejor.2008.03.044

Wright, P. D., & Mahar, S. (2013). Centralized nurse scheduling to simultaneously improve schedule cost and nurse satisfaction. *Omega, 41*(6), 1042-1052. doi:10.1016/j.omega.2012.08.004

Yildirim, D., & Aycan, Z. (2008). Nurses' work demands and work-family conflict: A questionnaire survey. *Int J Nurs Stud, 45*(9), 1366-1378. doi:10.1016/j.ijnurstu.2007.10.010

Chapter Six: Hospital Labor Planning and Allocation: A System Out of Balance

Abstract: Labor planning and allocation are critical success factors for hospitals in today's healthcare environment. Even though labor expenses can account for 50-60% of hospital total expenses, and hospitals typically are heavily focused on labor costs, there are still ample opportunities that exist. One possible explanation for the continued labor cost underperformance is the very structure in which labor planning and allocation occur. Hospitals are necessarily a "system of systems" consisting of multiple dynamic hierarchical subsystems that are interconnected and interdependent. Patient placement resources typically use a system view when managing patient flow through a hospital's various subsystems, but labor is typically managed from a departmental disconnected and siloed approach (especially forecasting and scheduling). We propose that the two conflicting structures of management processes prevents system balancing feedback loops from operating efficiently and subsequently contributing to poor outcomes. An elaborated Action Design Research project used a systems theory approach to redesign the structure, roles, and tasks of the labor planning and allocation cycle. New centralized roles dedicated to the tasks were developed to execute the functions of the cycle (forecasting, scheduling, and staffing). A new model of management was created to manage the functions from a system perspective and re-engage existing feedback loops. The artifacts created provide staff guidance regarding whom to interact with, what knowledge or resource to exchange, and what tasks to complete for each of the functions. Upon implementing this model in two hospitals across twenty-two departments, outcomes improved across the three functions. Additional benefits included giving hours back to department nursing leadership to focus on patients and employees after being freed from the back-office labor management functions.

Keywords: Action design research (ADR), Labor management, Scheduling, Staffing

Introduction

Hospitals in the United States are experiencing multiple pressures to reduce costs while at the same time improve the quality of care in a low margin environment. Both costs and quality of patient care have been demonstrated to be impacted by hospital staffing levels (Mark, Harless, McCue, & Xu, 2004; Needleman, Buerhaus, Mattke, Stewart, & Zelevinsky, 2002; Welton, 2011). Therefore, one of the keys to successful operational performance is the efficient and effective utilization of labor across the hospital organization. The challenge, however, typically rests with the organizational processes used to manage labor, which rarely develop on system structures. Instead, these processes and behaviors typically develop along organizational lines effectively creating segregated departmental silos. Department leaders tend to develop "self-protection" and "us vs. them" mentalities from a labor perspective focusing on individual departmental performance over considerations for organizational or service line needs. Often, this behavior can manifest itself in inter-departmental competition for resources effectively sub-optimizing the larger system staffing process. The behavior is not uncommon as noted in organizational behavior literature, the tendency for in-group favoritism and out-group animosity exists even when the groups can be dependent on each other (Ashforth & Mael, 2016). The subsequent culture evolves into a self-contained labor management entity, which then attempts to operate within the larger hospital system structure where interdependence and interoperability are required to facilitate smooth patient flow. It is the intersection of these two functions where we believe a conflict exists and sub-optimization negatively impacts the overall system. A vital connection between departments is severed or does not exist depriving the larger system of a critical source of information and resource exchange. We believe this disconnect contributes to poorer Labor Planning and Allocation (LPAC) outcomes that, in turn, contributing to a significant amount of time spent by nursing leaders performing these functions. Removing these functions from department nursing leadership responsibility will provide these same leaders more time to focus on patients and employees.

Subsystems disconnected from surrounding subsystems present a significant opportunity for transformative thinking to leverage a system view of the operation. There has been significant precedent work applying a system view to the patient side of the equation in the form of patient flow and patient throughput analysis, but the same rigor has not applied to the labor side of the equation. Academic studies and practitioners have relentlessly focused on the individual components of the LPAC cycle as demonstrated by the multiple literature reviews conducted in the areas of forecasting, scheduling and staffing (Castillo-Salazar, Landa-Silva, & Qu, 2016; Defraeye & Van Nieuwenhuyse, 2016; Park, Blegen, Spetz, Chapman, & De Groot, 2015; Van Den Bergh, Belien, De Bruecker, Demeulemeester, & De Boeck, 2013) These literature reviews cover projects and studies that have focused on analyzing and automating tasks to remove human work from the processes of forecasting, scheduling and staffing functions proposing various potential solutions to improve outcomes. The majority of these studies have applied their research to a narrow focus on a single labor management function within a similar narrowly focused environment (e.g., a single department, or type of department). We recognize a gap in current research where there has been a lack of investigation of the LPAC components as a lower layer subsystem operating within multiple higher-layer subsystems. While we see value in analyzing each function individually, we also see value in studying the collection of functions as they naturally exist within the larger complex hospital system. Therefore, this project takes a novel approach to the analysis of labor-management functions by applying a systems theory lens to view and rethink existing linkages between people, technology, and environment across the various subsystems that make up a hospital system.

Conceptual Framework

Since the 1950s, researchers have recognized the importance of viewing organizations as highly integrated collections of lower layer interdependent sub-systems with each of these elements being dynamic and interdependent to form a whole that is greater than the sum of its parts (Bertalanffy, 1950; Boulding, 1956; Meyer & O'Brien-Pallas, 2010). More recently, researchers have applied a system perspective to healthcare and specifically hospitals in multiple studies toward understanding the impact of systems on patient movement through the facility and the quality of care (Glover, Li, Naveh, & Gross, 2017; Marsilio, Torbica, & Villa, 2017; Tay, 2016). Patient movement and quality of care are only part of the equation, however. As patients move, resources must also move within the same system reacting to patient needs in higher layer balancing feedback loops. To fully understand the flows and interactions, one must view the labor side of the equation from the same system perspective as impacted by the surrounding environment.

Systems are visually understood better through stock and flow diagrams. In our case of department level labor management, we have a dynamic system where flows are both internally and exogenously controlled. As illustrated in Figure 15, patients and staff are two main flows of a simplified system diagram of a single departmental labor staffing subsystem. Stocks include patients currently in the department representing the workload and staff working in the department representing the inventory of labor. Flows are controlled via various balancing feedback loops (labeled "B" in the diagram) all constantly seeking a balance between the number of patients in the unit (workload) and the inventory of staff. Complexity exists in the fact that control on each of these flows can be internal or exogenous. Internal control exists when staff moves into or out of a department via controlled decisions (calling in additional staff at a premium, retaining staff on overtime, floating staff in from other departments, or utilization of contract labor, floating staff out to other departments or sending staff home). Exogenous control can simultaneously occur when staff unexpectedly call in sick leading to unbalanced conditions addressed through additional internally controlled responses. The same is true of patient flows. Controlled patient movement (internal control) occurs in response to balancing patient loads to staff (admitting or transferring more patients into a department or by discharging/transferring patients out of a department). Uncontrolled patient movement (exogenous control) occurs when patient care needs necessitate department movement regardless of staffing inventories.

An additional layer of complexity is that both patient and staff flow into or out of departments are subject to known and unknown constraints. For example, staff floats in or out of a department based on licensure and competency representing known constraints that limit staff mobility. Patients can be transferred in and out of a department based on care needs representing unknown constraints in advance and only known at the last minute.

The core opportunity addressed in this research is that a system perspective typically guides the management of patient flow, but a siloed, departmental perspective typically guides the management of labor planning and allocation. We posit that this lack of congruency challenges the balancing feedback loops of the system resulting in frequent and longer periods of imbalance causing cost and potential patient care implications in the form of understaffed or overstaffed departments. Applying a systems view of the associated forecasting, scheduling and staffing functions across multiple departments should allow the feedback loops to retain balance in the subsystem and allow for better outcomes. The method we believe that will allow for this system approach is the centralization of the labor allocation tasks into a single point dedicated to monitoring the various feedback loops and intervening with corrective actions when subsystems become out of balance. To develop this method, we considered various conceptual frameworks as foundations.

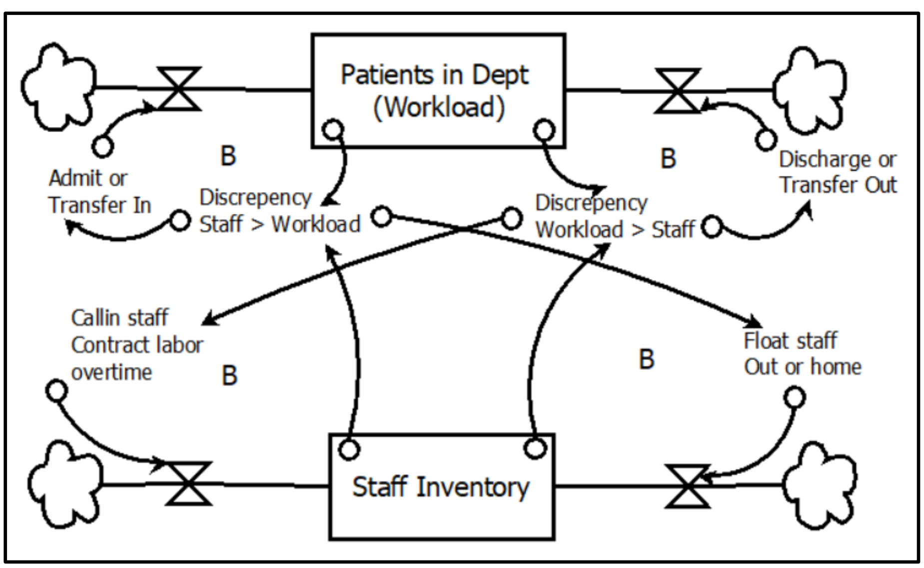


Figure 15 - Simplified Labor Allocation Stock and Flow Diagram

The sociotechnical system (STS) model was first discussed by Trist and Bamforth in the early 1950s to describe organizations based on their integration of the social and technological subsystems as well as the organizations' integration within the operating environment (Pasmore, Shani, Francis, & Haldeman, 1982). In this model, organizations consist of two main subsystems: a social subsystem including people, attitudes, individual relationships, and group relationships; and a technical subsystem including processes, structures, tools, and knowledge. (Shani, Grant, Krishnan, & Thompson, 1992). The overall success of an organization in producing a product or providing a service is related to the organization's ability to align and optimize these subsystems (Marsilio et al., 2017). This model is useful in studying the labor planning and allocation process in hospitals because the interaction of people and technology is a critical component of work in a healthcare environment and the human component of the work involved including knowledge sharing is a critical piece of process and structure design.

Taking the STS model one step further, human systems interaction (HSI) theory includes the basic assumption that most components of work include interaction between human beings, technology, and other aspects of the surrounding environment (Silva-Martinez, 2016). Closely aligned with this concept is Human Interaction Management (HIM) developed by Harrison-Broninski. HIM proposes that lost opportunities exist when technology is built around automated processes and serves as the sole mechanism for information exchange. Relying only on these automated processes ignores large parts of knowledge sharing and human interactions that can be significant components of task work. This type of environment funnels tasks into a specific definition and sequence that most likely does not match the definition and sequence in which tasks occur. In today's work environments, most workers do not follow pre-sequenced activities, but rather prioritize what to work on, seek knowledge to creatively solve problems and adapt to ever-changing criteria and environments that affect the parameters of the problems to be solved (Harrison-Broninski, 2005). Therefore, successful task completion is more dependent on the role participant's experience, ability to interact with others and known systems, and ability to seek out relevant information. STS is a stronger analytical framework than business process documentation in this environment because it includes the capability for work to be performed in different ways and sequences based on needs (Chisholm & Ziegenfuss, 1986).

Additionally, HIM includes concepts to support human interaction and learning that existing business process management neglects. (Harrison-Broninski, 2005; Tarpey, 2017). The context of this research is the integration of human beings, processes, knowledge, and technology across departments and service lines of a hospital to create a model that provides role participants with guidance regarding whom to interact with, what knowledge or resource to exchange, and what tasks to complete. HSI and HIM provide a framework for understanding the integration of the social subsystem (humans) and the technical subsystem (process, knowledge, and technology) within the labor planning and allocation space. To successfully develop a new model of centralized labor allocation, we needed to consider both the technical and social subsystems in which this new model will operate in as well as the many human-human and human-systems interactions involved with forecasting, scheduling and staffing.

Method

This project was an extension of two prior investigative research projects. The first research project employed an elaborated Action Design Research (eADR) methodology to produce measurement artifacts useful in measuring and assessing both intermediate and end outcomes from the labor planning and allocation cycle. The metrics were developed and tested by the embedded researcher and practitioners over several iterations of the eADR define, build, evaluate, learn, reflect, and redefine cycle. The methodology allowed the team to validate metric artifacts that met four key criteria: adoption, time-series comparisons, organizational level comparisons, and system perspective (detailed in Chapter 1). Refer to Table 25 for definitions of these requirements.

Table 25 – LPAC Performance Metric Acceptance Criteria

Acceptance Criteria	Description
Adoption	Metric must be easy enough to understand to allow practitioners to assess performance and determine either corrective or confirming actions
Time Series Comparison	Metric must allow for valid comparisons across multiple time series
Organizational Level Comparisons	Metric must allow for valid comparisons across multiple organizational levels
System Perspective	Metric must allow for the ability for aggregation (rolled up) to measure the combined efforts of multiple organizational structures

The second research project also employed the eADR methodology embedding a researcher within an organization to co-design and co-create a new model of labor planning and allocation taking a system view of forecasting, staffing, and scheduling in a limited test environment. The main objective of the project was to develop a new model for managing and executing these functions to achieve better outcomes. The model leveraged the interdependencies of organizational departments to facilitate communication, knowledge sharing, and information flow to optimize system performance overcoming sub-optimization of the “every department for themselves” approach (Tarpey, 2017 & Chapter 2). The current research builds upon the artifacts developed in the prior two projects and seeks to generalize these artifacts across a larger test domain. We extended artifact development and refinement to demonstrate their ability to be applied to larger environments concentrating on how the Systemic LPAC can operate in a typical hospital and how a system approach driven by patient flow and not limited by departmental structures leads to greater optimization of the balance between nursing labor, nursing satisfaction, and patient outcomes. We sought to understand the answers to these questions from the perspective of model refinement, implementation of a full-scale central labor planning office and analysis of performance outcomes. The artifacts from the prior two projects were applied to a larger test domain consisting of two hospitals and studied for a longer period to understand the potential for sustainability.

Our research employed a four-stage elaborated Action Design Science Research (eADR) model (Mullarkey & Hevner, 2015) to investigate alternative designs of structures, roles, and processes of the labor planning and allocation cycle accommodating a system perspective while optimizing integration between the social and technical subsystems. The researcher embedded within the organization to design, test and implement a centralization of forecasting, scheduling, and staffing across two hospitals that had previously performed the tasks from the individual departmental perspective approach. The methodology consisted of multiple iterations of the define, build, evaluate, learn, reflect, and redefine cycle working with role participants to verify and validate each existing and new artifact. Data collection was accomplished via multiple sources during the project in the form of notes, interviews, observations, and report analysis over the course of twelve months. Project outcomes extended beyond artifact creation to include the transition of labor allocation cycle functions from the individual facility departments to a fully functional, centralized service center. Performance of the center was assessed and compared back to pre-transition performance. Additionally, the research documented several other identified benefits resulting from the utilization of the system perspective in this environment.

The Hospital as a System

Viewing the hospital as a system reveals complexity. The overall higher layer system consists of several service lines, which in turn consist of multiple individual departments that are interdependent. While each of these service lines is considered separate entities, patients flow between departments within a service line and across different service lines during their treatment and care. The pathway through these subsystems is not consistent, but unique to each patient based on illness, acuity, and care needs. The concept of a “system of systems” presented by

Silva-Martinez is particularly attractive for conceptualizing the complexity (Silva-Martinez, 2016). Silva-Martinez visualized complex systems in the form of sets of subsystems interfaced together which is descriptive of the hospital system (refer to Figure 16). The diagram represents the hospital inpatient system (first-layer system) consisting of hierarchical subsystems. The second-layer subsystem is made up of the different service lines within inpatient services. The third-layer subsystem includes the individual departments, and the fourth-layer subsystem is the labor forecasting, scheduling, and staffing functions. Patients and resources flow between these subsystems including multiple integrations of technological (e.g., information, data, knowledge) and social (e.g., relationships, attitudes) components.

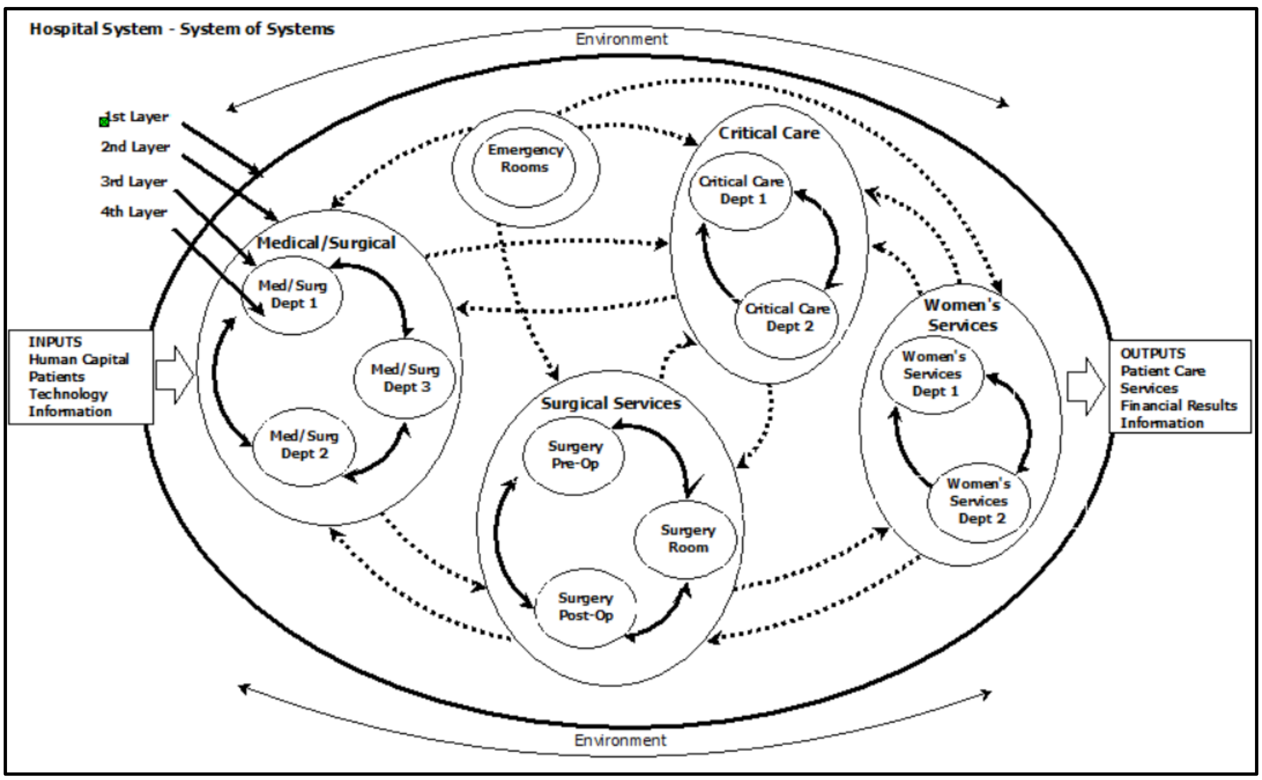


Figure 16 - Hospital Organizational Structure System of System Complexity

This visualization of the system supports a clearer understanding of the challenges presented if a department is operated as an unconnected workgroup with no integration linkages to like departments within a service line subsystem or to other higher layer subsystems. When a department operates independently from a labor planning, and allocation perspective, technological and social integration connections are severed, however patient flow continues between departments as illustrated in Figure 17. Isolating one service line or a department within a service line creates imbalances in knowledge and resources and interrupts necessary feedback loops that are critical to keeping the hierarchical subsystems and in turn overall system in balance.

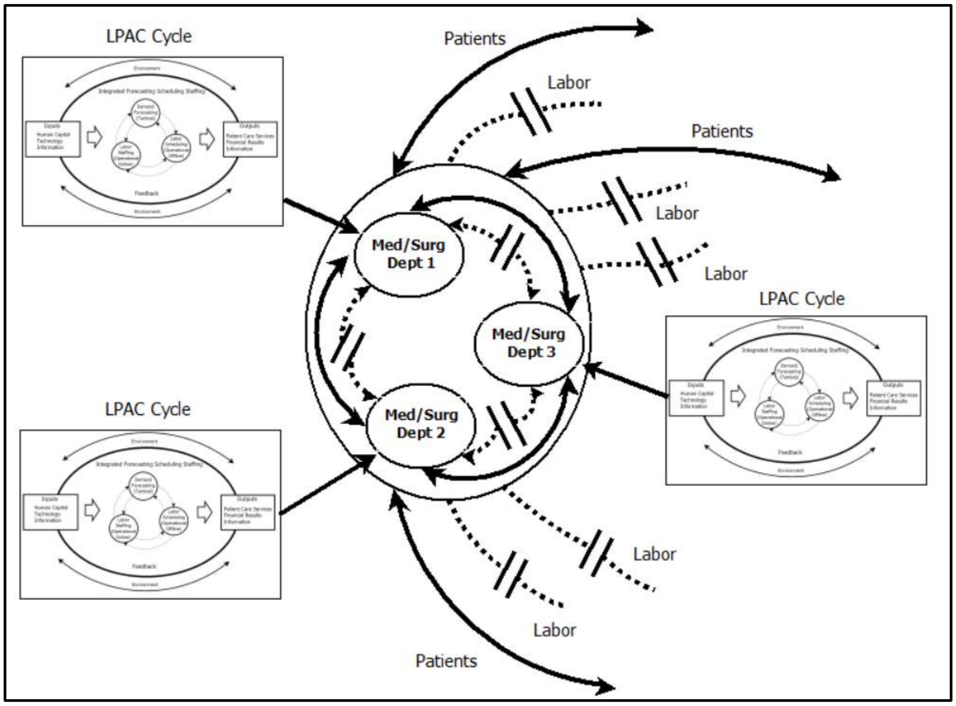


Figure 17 - Operating Departments or Service Line Isolation from Higher Layer Systems

Labor Forecasting, Scheduling, and Staffing Subsystem

This project's main focus is labor planning and allocation specifically, forecasting, scheduling, and staffing labor in a manner that adequately plans for, schedules, and staffs employees to cover patient needs. These tasks represent the lower, fourth-layer subsystem in our model. In this subsystem, multiple inputs transformed during the tasks of forecasting, scheduling, and staffing produce multiple outputs (refer to Figure 18). When a hospital takes a non-system or individual, siloed approach, these subsystems typically operate at the departmental level, executed by departmental leadership roles such as directors, managers, and charge nurses, which was the pre-implementation case in the hospitals involved in this study. The departments were, for the most part, self-contained during the execution of these tasks with very little consideration given for departmental connections and interdependencies. Communication between departments was nonexistent during the forecasting and scheduling phases and limited to house supervisor coordination of staff at the last minute just before and during the shift while staffing.

The interdependencies within this subsystem are essential to recognize. The three phases are cyclical repeating on a regular cadence to forecast the workload, create the labor schedule based on the forecast, and staff the department based on the schedule and patient/departmental needs. The activities in each of these phases have the potential to impact and alter activities in a downstream phase as output from one phase is consumed in the next phase as input or consumed in an upstream phase as feedback. (Hulshof, Kortbeek, Boucherie, Hans, & Bakker, 2012; Tarpey, 2018). The potential interactions are significant because they can be determinants in both the quality of outputs and the amount of effort to achieve quality outputs. For example, an inaccurate workload forecast can be either used or ignored in the scheduling phase. If the scheduler follows the forecast, then the effort to create the schedule will be lower than if the scheduler ignores the forecast and creates an overriding forecast that he/she believes to be more accurate. The quality of the schedule, however, will be lower if the scheduler follows a forecast that proves to be inaccurate. Likewise, an ineffective or inefficient schedule can be overcome in the staffing phase but will require more effort to find or reallocate staff at the last minute. Essentially, higher effort in subsequent phases can compensate for poor outcomes in a previous phase (refer to Figure 19).

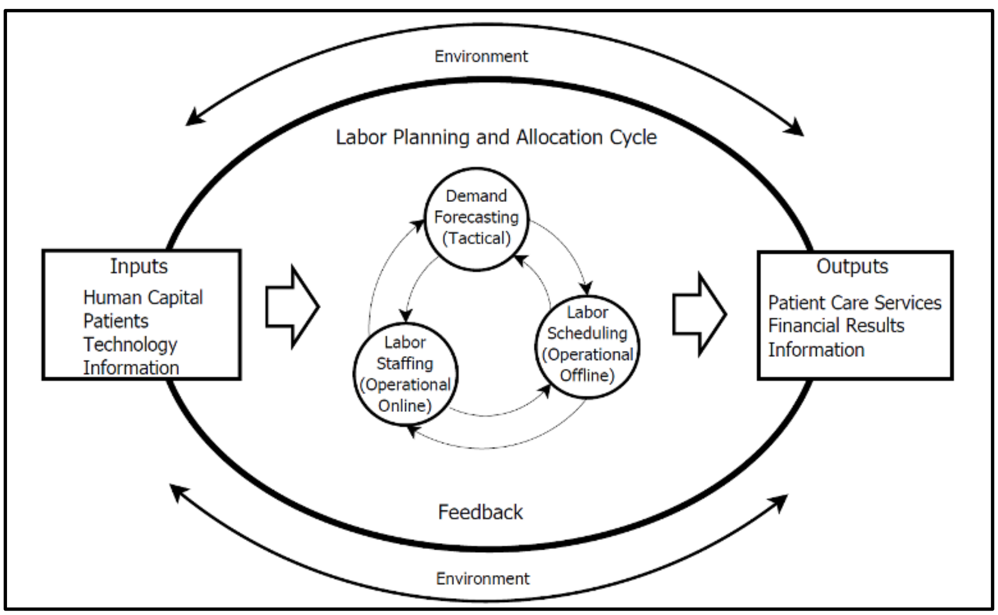


Figure 18 - Labor Planning and Allocation Cycle (LPAC)

Beyond understanding the interactions and dependencies of the LPAC Cycle phases, we needed the ability to measure the performance of each phase. The performance assessment could then be used to determine if the centralization of tasks under the new management model had any impact on the quality of the outcomes associated with each phase and how the quality impacted the work of the associated job roles. We used the performance metrics outlined in Tarpey 2018 to measure pre-implementation and post-implementation outcomes: F-SbMAE (forecasting accuracy), Completeness, Professional Completeness, Support Completeness, Commitment, Healthiness, and Preferences (scheduling) and S-SbMAE (staffing accuracy) (Tarpey, 2018). Over multiple LPAC cycles, these metrics were calculated to assess the quality of outcomes and accordingly track both confirming and correcting actions taken at the various intervention points. Upon completion of each of the cycles, the team assessed performance and analyzed from the perspective of the RADs and IMCs to determine if the system perspective was providing value and if any model adjustments were needed. This methodology provided the iterative define, build, evaluate, learn, reflect, re-define process to adequately define and test the artifacts created.

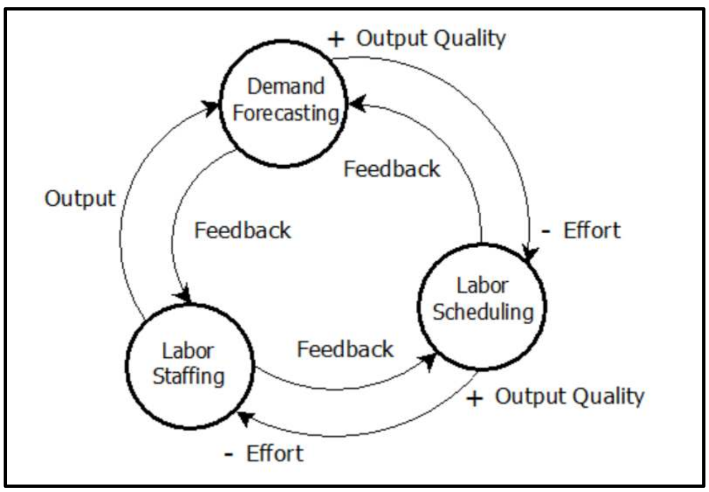


Figure 19 - LPAC Quality & Effort Relationships

LPAC Management Model Development

Glover et al. described two categories of the technical subsystem: formal and informal integration. Formal integration includes clearly defined job descriptions, authority structures, responsibilities, accountability, information systems, protocols, procedures, and workflow systems all making up the formal practices of how tasks get completed and knowledge is processed. Informal integration includes casual information exchanges by any means sharing knowledge and information (Glover et al., 2017). Efforts were made in our model design to accommodate and include aspects of both formal and informal integration mechanisms to provide role participants guidance on where critical knowledge can be found to execute given tasks. The artifacts developed were painstakingly refined over multiple iterations as each service line of both hospitals moved into the central labor planning office adapting to the new service line specific or hospital-specific nuances presented in forecasting, scheduling, and staffing. The artifacts, however, were kept at a high level with the intent of documenting and providing standardized guidance on where solutions exist rather than becoming detailed process map documentation attempting to provide pre-determined but most likely non-optimal solutions. This aspect of the documentation was meant to coincide with the fluid and dynamic environments of each of the functions giving role participants necessary guidance, but also the flexibility to be creative and develop custom solutions to challenges presented. In this manner, generalized artifacts were developed to apply to multiple hospitals without the need for constant modifications. This strategy also created artifacts that were obsolescence resistant. The artifacts did not need modification with every hospital operational change since the business rules and knowledge retained in systems were accessible by role participants when needed, but were not hardwired parts of the process documentation.

The LPAC model starting point for this research was the artifacts generated from the prior LPAC management model development project (Tarpey, 2017 & Chapter 2). The current project continued with the use of the Role Activity Diagram (RAD) methodology for the documentation of roles, tasks and knowledge sharing due to the tool's intuitive nature for modeling processes and interactions (Ould, 1995; Phalp, Henderson, Walters, & Abeysinghe, 1998). These diagrams proved useful in explaining work streams to role participants working in the newly developed central planning office roles as well as departmental and facility leadership at the hospitals. The documentation was extended to include Interaction Model Cards (IMC) to visualize and understand human-human and human-system interaction critical to the successful outcomes of the work (Seo, Yoon, Lee, & Kim, 2011). Figures 20 and 21 provide an example of a RAD for forecasting and scheduling along with an example of an associated IMC for one of the processes. These example artifacts represent the definitions of the new model of LFSS cycle management containing the guidance for central planning staff to understand what knowledge and information are needed to complete tasks and where the knowledge and information resides (what interactions are necessary).

The RAD example presented in Figure 20 shows the various work roles across the top and details the functions in which each role is involved. The diagram also illustrates examples of knowledge and information flow further documented within the IMCs. Figure 21 provides the IMC associated with one of the processes (P1) in Figure 20. The IMC details the goal, participants, responsible role, interactions (human-human and human-system) required, touchpoints required, and knowledge shared and used in the process. The interactions documented during this process were critical in designing both the job roles and the work structures of the central planning office. As the forecasting, scheduling and staffing functions for each service line of the hospitals moved into the central planning office; the management model was reviewed, analyzed, and refined based on prior experiences and learning. The resulting process was a continuous verification effort reinforcing the artifacts over multiple iterations.

An essential aspect of the artifacts is that they are not departmentally based and represent work performed in the overall system of subsystems. While information or knowledge stores may vary from subsystem to subsystem, the artifacts do not change providing unique flexibility and obsolescence resistance over standard process mapping documents. Resources performing the LPAC tasks can understand the existing relationships and are guided to seek information from all necessary sources across the system to complete tasks. The simplified high-level causal loop diagram in Figure 22 was used to illustrate the core conceptual model of the centralized planning office functions and interdependencies detailing two service line subsystems and two departmental subsystems as an example. The diagram shows how the various LPAC subsystems (forecasting, scheduling, and staffing) interact with the departmental, service line subsystems and the hospital system including the external environment. Labor planning and allocation subsystems operate within the individual departments but are interdependent with the other departments within the service line subsystem from a perspective of patient and staff movement in the scheduling and staffing functions. The service line subsystems also have the potential for interdependency with the other service line subsystems from the perspective of patients, which is likely and from the perspective of staff assuming appropriate licensures, credentials, and competencies. Additionally, events within the hospital system and the external environment such as local incidents (e.g., large freeway accident) or regional incidents (e.g., flu outbreak, hurricane) can impact any of the given subsystems.

These interdependencies were used to design the formal integration aspects of the central planning center. While not uncommon for per shift staffing to already accommodate a higher layer systems perspective in hospitals, forecasting, and scheduling functions had to be redesigned to be monitored and assessed from the service line subsystem level down to the department level. Each of the roles and functions was crafted to facilitate human-human and human-systems interactions including touchpoints and potential intervention points. The RADs, IMCs, and system diagram artifacts provided guidance concerning work tasks, role responsibilities, and interaction points for each of the various functions whether human or system in nature.

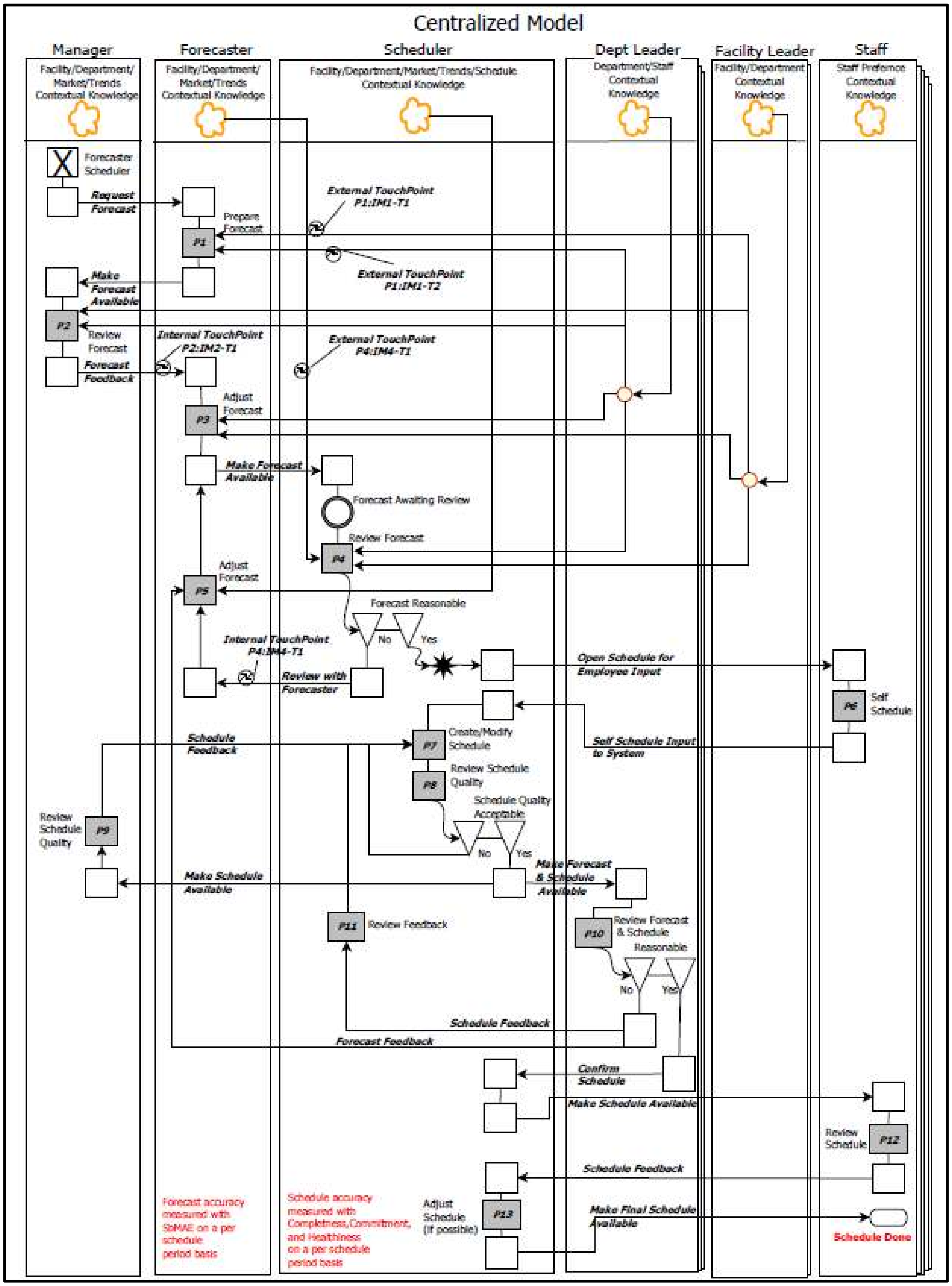
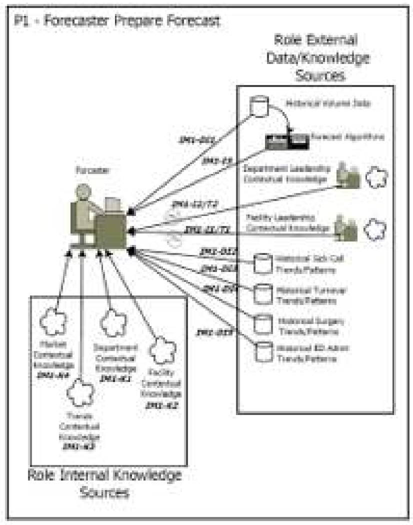


Figure 20 – Example Role Activity Diagram – Forecasting and Scheduling

Process 1 – Prepare Forecast Interactions (P1)



Interaction Model Card (IM1):

Goal: Prepare labor forecast for future labor schedule period

Participants: Forecaster, Department Leaders, Facility Leaders

Responsible Role: Forecaster

Interactions:
(IM1-I1) Forecaster understands departmental conditions from department leaders
(IM1-I2) Forecaster understands facility/departmental conditions from facility leaders
(IM1-I3) Forecaster reviews historical data and runs forecast models via applications

Touchpoints:
(IM1-T1) Forecaster – Department Leader (Forecaster leads)
(IM1-T2) Forecaster – Facility Leader (Forecaster leads)

Knowledge:
(IM1-K1) Forecaster – contextual knowledge concerning departmental conditions
(IM1-K2) Forecaster – contextual knowledge concerning facility conditions
(IM1-K3) Forecaster – contextual knowledge current volume trends
(IM1-K4) Forecaster – contextual knowledge concerning market trends

Input Data:
(IM1-DI1) Historical volume data
(IM1-DI2) Sick call trend/pattern analysis
(IM1-DI3) Turnover trend/pattern analysis
(IM1-DI4) Surgery schedule trend/pattern analysis
(IM1-DI5) Emergency admission to unit trend/pattern analysis

Output Data:
(IM1-DO1) Historical volume data

Figure 21 - Example Interaction Management Card – Process P1

Results Summary

Over the course of the twelve-month project, twenty-three nursing departments from three service lines of the two hospitals phased into the new model (medical/surgical, critical care, and emergency services). The approach taken was to compare pre-implementation outcome metrics for the individual departments with post-implementation outcome metrics for the identified subsystems. Hospital A consisted of five medical/surgical, two critical care and one emergency services department. Hospital B consisted of twelve medical/surgical, three critical care, and one

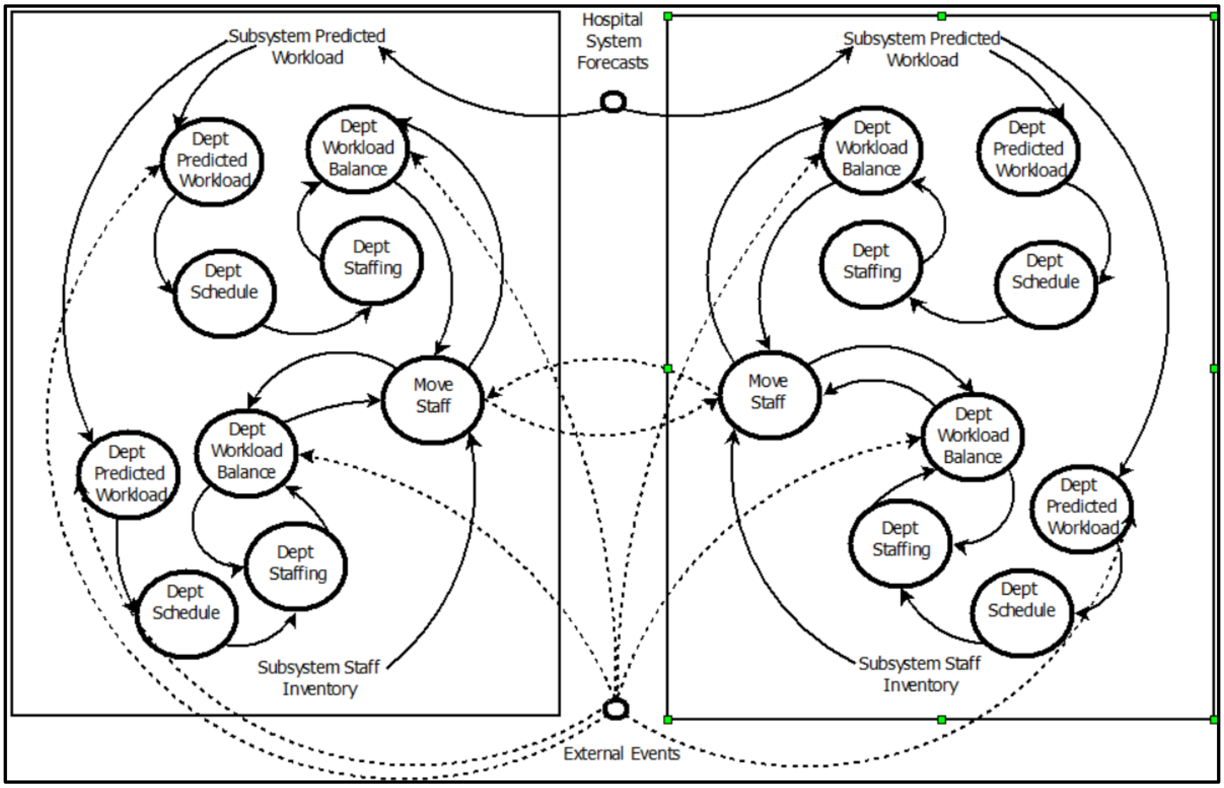


Figure 22 - Simplified Causal Loop Diagram (2 Subsystems)

emergency services department. The pre-implementation outcome measurement time horizon consisted of a minimum of three months before implementation, and the post-implementation outcome measurement time horizon consisted of a minimum of three months post-implementation. We analyzed results in three ways. First, we looked at simple counts of department performance pre and post-implementation concerning improved outcome metrics. Second, we calculated a difference score for each case as the pre-score minus the post-score such that positive numbers reflect cases in which the new model outperformed the old model (assigned 1), zero reflects cases in which the model performed equally to the old model (assigned 0), and negative numbers reflect cases in which the new model did not outperform the old model (assigned -1). We then subjected these scores to a single-sample t-test to determine if the new model resulted in significant improvements. Third, we compared performance metrics from the individual departments pre-implementation to service line metric results post-implementation to determine if the new model resulted in outcome improvement across the service line system.

It is critical to determine where in the cycle a systemic view of performance provides more value than a departmental view. In the workload forecasting phase, we found that third-layer (departmental) forecasts were more useful as a starting point subsequently aggregated during the scheduling phase for ensuring a complete labor schedule was created covering the second-layer (service line) labor needs. Forecasting a pre-aggregated volume number at the service line level did not provide the discrete values needed to schedule labor at the department level but did provide an overall confirmation factor when comparing against historical service line loads. In the scheduling phase, we found that the more useful indicator resided at the service line level ensuring enough labor scheduled within this subsystem with less relevance placed on which specific department the labor was scheduled (i.e., scheduling labor into home departments). The common practice of floating staff across departments within a service line subject to appropriate constraints supported this finding. In the staffing phase, we found that the department subsystem metrics were critical for ensuring that staffing met patient needs for each department.

Workload Forecasting

The LPAC is iterative as illustrated in figure 5.3. Outputs from one phase become inputs to the next phase. Therefore, it is critical to analyze performance at each phase of the cycle to determine if correcting action is needed and if so what action to take. Workload forecasting is the beginning point of the cycle. The expected workload is the foundation for building labor schedules since the expected workload per shift determines the number of staff per skill set to be scheduled. An accurate forecast will allow for a labor schedule that better prepares the department for the expected workload. The measurement used to assess the accuracy of the forecast is the Forecast Staffing Bin Mean Average Error (F-SbMAE) (Tarpey, 2018 & Chapter 1). The "Staffing Bin" refers to the staffing categories on a nursing department staffing grid which determines the number of each skill set required to be working in the department. For example, a department with a 1:2 nurse to patient ratio requires one nurse for either 1 or 2 patients, two nurses for either 3 or 4 patients and so on (refer to Table 26 for the details of the example). The measurement represents the mean average error between the forecasted staffing bin and the actual staffing bin required at the time of the shift (Tarpey, 2018 & Chapter 1). An F-SbMAE value of 1.00 indicates that the workload forecast was on the average one staffing grid bin away from the actual staffing grid bin required, which means a change in labor must occur (either adding or removing staff). Continuing the example, an actual value of seven patients requires labor corresponding to bin number 4 on the staffing grid. A forecast of five patients represents labor from bin number 3 which is one labor bin away from the forecasted labor bin. The forecast error in this example is (4-3) = 1.0 for one shift. The F-SbMAE for an organizational structure for a schedule period is the average of the errors for the given time horizon. An F-SbMAE value of 0.00 represents a perfect forecast. Refer to Chapter 2 for calculation details.

Table 26 – Example Staffing Grid

Patients Range	RN	Tech	Bin Number
0 – 2	1	0	1
3 – 4	2	0	2
5 – 6	3	1	3
7 – 8	4	1	4
9 – 10	5	1	5
11 - 12	6	2	6
13 – 14	7	2	7

In the pre-implementation model, forecasting was performed by the Chief Financial Officer (CFO) at one hospital and by each department director at the second hospital resulting in limited interactions and subsequently limited input information. Observation notes demonstrated that there was very little interaction across departments in either hospital. Even in the case where the CFO created the forecast, there was little input from the department leadership and equally little consideration for forecasts across service lines. The new model called for a different approach. The new central forecasting role generated forecasts for the individual departments based on information and knowledge gathered from multiple sources across the service line subsystems as documented in the RAD. The forecasts were then rolled up at the service line level to gain an expectation of workload across the subsystem and the impact of the forecast on other service line subsystems. The employee performing the forecast sought out information from multiple system and human sources as indicated in the IMCs resulting in more knowledge engaged in the process and a more focused effort. Formal and informal integration components were leveraged to gather input. Results demonstrated improvement.

We tested potential workload forecasting improvements in two ways. First, the new procedure significantly outperformed the old procedure in simple counts in 20 of 24 cases (85.71% improved; $\chi^2(1)=10.71$; $p=.0011$). At Hospital A, four of the five medical/surgical departments averaged a higher forecasting accuracy in the post-implementation period. Both critical care departments averaged higher forecasting accuracies as well as the emergency services department. There was one outlier department (Department 3) where the forecast accuracy was lower over the post-implementation time horizon. In this case, the department went through a high volume variability state in the post-implementation period causing forecasting to be significantly more difficult. The number of patients varied from a low of five to a high of fifteen regularly oscillating between the two with no discernable pattern. One of the outcomes of the new model was that the volume variability was recognized and regularly discussed with the scheduling and staffing analysts to adapt while performing downstream functions. The increase in communication prevented a negative impact on staffing as indicated by staffing performance. Essentially, the model was able to accommodate the more variable volume fluctuations in a better manner as demonstrated by staffing accuracy.

Hospital B has similar results. Ten of the twelve medical-surgical and all of the critical care and emergency departments experienced forecast accuracy improvement. In the two departments where a decline in accuracy occurred, both had intermittent unforeseen periods of closure due to low patient volumes. The central planning office potentially should have foreseen these closures. When the department closed without the central planning office expecting the event, the team met and identified additional touchpoints for communication between the facility and the planning office regarding unit closures on a daily basis. These new intervention points and touchpoints became part of the new model. This occurrence was central to the flexibility and adaptability of the new model.

We also calculated a difference score for each case as the pre-score minus the post-score such that positive numbers reflect cases in which the new procedure outperformed the old procedure. We then subjected these scores to a single-sample t-test which found that the new procedure was significantly more accurate than the old (M=0.20; $t(21)=3.13$; $p=.0053$). Overall, workload forecasting accuracy results under the new model outperformed results under the prior model (refer to Table 27). Workload forecast accuracy also improved at the service line level. In Hospital A, the Medical/Surgical service line F-SbMAE metric improved from 0.99 to 0.91 while in the Critical Care service line F-SbMAE metric improved from 1.30 to 1.16. In Hospital B, the Medical/Surgical service line F-SbMAE metric improved from 0.48 to 0.32 while in the Critical Care service line F-SbMAE metric improved from 1.03 to 0.46.

Scheduling

Labor scheduling is the next phase of the cycle. In this stage, six metrics were used to assess the quality of the labor schedule. The metrics included (Tarpey, 2018; Tarpey & Nelson, 2009):

- *Completeness* – measures schedule effectiveness meeting workload demand as measured by the percentage of schedule slots filled but not overfilled
- *Professional Completeness* – same as Completeness, but limited to nursing staff

- *Support Completeness* – same as Completeness, but limited to support staff
- *Commitments* – measures percent of staff scheduled to full allocation or commitment
- *Healthiness* – measures the percent of staff scheduled to healthy schedules (absence of unhealthy shift patterns)
- *Preferences Honored* – measures the percentage of staff schedule preferences honored in the final pre-worked schedule

Table 27 – Workload Forecasting Results

Hospital	Service Line	Dept	Pre-State Average F-SbMAE	Post-State Average F-SbMAE	Diff	Binary 1=improved 0= not improved
A	Medical/Surgical	1	1.09	0.83	0.260	1
A	Medical/Surgical	2	1.03	1.01	0.020	1
A	Medical/Surgical	3	0.50	0.79	-0.290	-1
A	Medical/Surgical	4	0.92	0.76	0.160	1
A	Medical/Surgical	5	1.42	1.19	0.230	1
A	Critical Care	6	0.88	0.74	0.140	1
A	Critical Care	7	1.73	1.59	0.140	1
A	Emergency	8	1.47	0.98	0.490	1
B	Medical/Surgical	1	0.47	0.43	0.040	1
B	Medical/Surgical	2	0.92	0.19	0.730	1
B	Medical/Surgical	3	0.72	0.54	0.180	1
B	Medical/Surgical	4	0.12	0.16	-0.040	-1
B	Medical/Surgical	5	0.23	0.13	0.100	1
B	Medical/Surgical	6	0.49	0.47	0.020	1
B	Medical/Surgical	7	0.26	0.23	0.030	1
B	Medical/Surgical	8	0.64	0.31	0.330	1
B	Medical/Surgical	9	0.77	0.36	0.410	1
B	Medical/Surgical	10	0.31	0.35	-0.040	-1
B	Medical/Surgical	11	0.36	0.32	0.040	1
B	Medical/Surgical	12	0.00	0.00	0.000	0
B	Critical Care	13	0.54	0.17	0.370	1
B	Critical Care	14	1.91	0.82	1.090	1
B	Critical Care	15	0.65	0.40	0.250	1
B	Emergency	16	1.01	0.87	0.140	1

In the scheduling phase, it is the service line subsystem that takes a higher priority. The new LPAC model indicates a preference for ensuring enough staff planned across the service line with the expectation and understanding that staff move between like departments subject to constraints. Therefore, while in the old model, departments were either over-scheduled or under-scheduled based on the director's ability to complete a schedule with given resources, in the new model schedules were balanced across the service line to compensate for overages and underages in each department. The difference in methodology presented a challenge with outcome comparisons. We compared individual department schedule quality scores pre-implementation and post-implementation, but the better assessment of performance came from the comparison of schedule quality scores at the service level subsystem. Table 28 provides a summary of Schedule Quality results comparing pre-state department performance against service line performance.

The comparisons of individual department pre-implementation and post-implementation results did not show significant improvement in the simple counts or single sample t-test analysis. We were not overly surprised by these results for the three completeness metrics since during the scheduling phase more emphasis was placed on balancing schedules across the higher layer service line subsystem with little focus on individual department metric improvement. Overscheduled departments might be left overscheduled if there was a verifiable need within the service line due to a corresponding department being under scheduled due to a lack of resources (e.g., vacancies). In this example, the individual department scores would be poor, but the service line score would be good, which is more useful. The staffing resource can be confident that resources are available within the service line to cover all needs. The commitment and healthiness metric also did not show significant improvement within individual departments, but again we were not overly surprised. Performance in these two metrics was strong before implementation and did not suffer significant decay in performance under the new model. We viewed this as a positive result. Interestingly, the efforts of the central resources may have decreased performance in the Healthiness metrics. The scheduling resources worked to fill holes in the schedule via more communication to staff resources who were willing to volunteer overtime. While filling these holes improved the Completeness of the schedule, there was an increase in unhealthy shift patterns due to the voluntary overtime. Volunteers had to be individually analyzed to prevent overly unhealthy patterns (i.e., limitation to volunteer overtime). Individual department metric comparisons follow:

- Total Completeness:
    - Simple Counts: (63.64% improved; $\chi^2(2)=11.55$; $p=.0031$)
    - Single sample t-test: (M=0.05; t(24)=1.93; $p=.0674$)
- Professional Completeness:
    - Simple Counts: (45.45% improved; $\chi^2(2)=8.27$; $p=.0160$)

- Single sample t-test: (M=0.01; t(22)=0.71; p=.4871)

- Support Completeness:
  - Simple Counts: (68.18% improved; $\chi^2(2)=13.72$; p=.0010)
  - Single sample t-test: (M=0.05; t(22)=1.56; p=.1339)
- Commitment:
  - Simple Counts: (50.00% improved; $\chi^2(2)=0.00$; p=1.0000)
  - Single sample t-test: (M=0.02 ; t(22)=2.65; p=.0149)
- Healthiness:
  - Simple Counts: (50.00% improved; $\chi^2(2)=0.00$; p=1.0000)
  - Single sample t-test: (M=0.012; t(22)=0.30; p=.7654)

The service line metrics exhibited more significant improvement. Hospital A medical/surgical departments schedule quality increased under the new model when comparing the average schedule quality metrics per department pre-implementation versus the schedule quality metric for the service line in the new model. The improvement occurred across four of the five quality dimensions. The only exception was Commitments which remained at 100% in the old and new models. In critical care, the results were less pronounced. Schedule quality increased slightly in the Completeness and Professional Completeness metrics. It remained the same for Support Completeness metrics and increased slightly in the Commitments and Healthiness metrics. Several factors identified might have contributed to these results. The service line is made up of just two departments limiting the pool of labor for covering underages in the schedule. Both departments, during the post-implementation time frame, had multiple nurse vacancies and therefore where short nurses to complete the schedule. There were also limited contract labor opportunities in this market. Schedules were completed with high amounts of volunteer pre-scheduled overtime to fill in open needs, but in both departments, the volunteer overtime was not enough to compensate for the number of unfilled needs. The resulting schedules during the post-implementation period were unscheduled due to a lack of resources. The 100% Commitment metric indicates that every staff member in both units had enough hours scheduled to meet hours available. Even though there were more nurse vacancies in the service line and schedules were under scheduled, staffing performance was no worse than before implementation. The coordination between the scheduling and staffing resources within the central planning office identified the service line exception for intervening actions to occur, such as contacting staff to request more voluntary overtime and pre-requesting contract labor resources from vendors.

Table 28 – Scheduling Results

Service Line	Hospital	Metric	Pre-Implementation Department Average	Post-Implementation Service Line Average
Med/Surg	A	Completeness	80.80%	85.92%
Med/Surg	A	Professional Completeness	84.92%	92.54%
Med/Surg	A	Support Completeness	59.24%	76.38%
Med/Surg	A	Commitments	100.00%	100.00%
Med/Surg	A	Healthiness	93.60%	94.32%
Critical Care	A	Completeness	85.95%	86.40%
Critical Care	A	Professional Completeness	85.66%	86.07%
Critical Care	A	Support Completeness	85.97%	85.97%
Critical Care	A	Commitments	99.92%	100.00%
Critical Care	A	Healthiness	91.11%	91.77%
Emergency	A	Completeness	73.15%	78.37%
Emergency	A	Professional Completeness	73.09%	80.86%
Emergency	A	Support Completeness	81.26%	83.88%
Emergency	A	Commitments	97.97%	100.00%
Emergency	A	Healthiness	86.33%	90.79%
All	A	Preferences Honored	N/A	89.72%
Med/Surg	B	Completeness	75.34%	82.19%
Med/Surg	B	Professional Completeness	83.09%	87.41%
Med/Surg	B	Support Completeness	58.59%	66.31%
Med/Surg	B	Commitments	98.71%	99.27%
Med/Surg	B	Healthiness	94.17%	93.73%
Critical Care	B	Completeness	71.72%	85.40%
Critical Care	B	Professional Completeness	71.72%	85.07%
Critical Care	B	Support Completeness	100.00%	100.00%
Critical Care	B	Commitments	86.21%	94.68%
Critical Care	B	Healthiness	84.37%	84.76%
Emergency	B	Completeness	80.54%	86.87%
Emergency	B	Professional Completeness	80.92%	83.54%
Emergency	B	Support Completeness	69.52%	79.25%
Emergency	B	Commitments	80.19%	100.00%
Emergency	B	Completeness	84.34%	86.08%
All	B	Preferences Honored	N/A	93.23%

In Hospital B, schedule quality improved in the medical-surgical service line in all metrics except for Healthiness. This metric includes voluntary overtime, which increased during the post-implementation period. The central planning center provided more opportunities and communication regarding open shifts to hospital staff resulting in higher rates of volunteer overtime providing support to filling the schedules, but also causing more unhealthy shift patterns even though they were voluntary as in Hospital A. The critical care service line performance improved in all dimensions of schedule quality. The commitment score improved due to a larger effort performed by the central planning center in getting PRN employees to pre-schedule shifts to fulfill their agreements.

Preferences Honored was measured in the post-state, but not available from the pre-state period. Therefore, comparisons are not possible regarding the increase or decrease in preferences honored. Hospital and departmental leadership, however, felt the results were valuable as an employee satisfaction communication tool to staff (near 90% average for Hospital A, and 93% average for Hospital B).

Staffing

Staffing is the final phase of the LPAC. Several internal and external variables impact staffing in addition to the workload forecast and the labor schedule. This phase is the phase that ensures the right staff is in the departments at the right times resulting in critical work to contain costs. The central planning center staffing analysts were on duty 24x7x365 ensuring that hospital department staff matches the workload needs of the departments. This 24-hour focus provides real-time review and communication with the hospital across both formal and informal integration mechanisms continually adjusting the staffing quantities and skill mixes. The same concept was used to assess staffing performance as with volume forecasting concerning the staffing grid. The definition of performance in staffing is staffing the correct number of each skill set as indicated on the staffing grid for the given patient load. A Staffing Bin Mean Average Error (S-SbMAE) measurement compares the staffing grid bin number of the actual staff on the floor to the staffing grid bin number that corresponds to the actual number of patients. Revisiting the example from the workload forecasting section, if the number of patients in the department is ten corresponding to a bin number of 5 and the number of nurses working in the department is four, corresponding to a bin number of 4, then there is a staffing miss. The miss equates to one labor bin (5-4) indicating either a shortage or overage of staff. The S-SbMAE over a time horizon is the average of the errors for each data point. When this value is greater than 1.00, there is cause for corrective action as the unit is consistently working either short staffed or overstaffed for the period. (Tarpey, 2018 & Chapter 2).

Overall, staffing accuracy results under the new model outperformed results under the prior model (refer to Table 29). 21 of 24 departments performed better in the new model (86.36% improved; $\chi^2(1)=11.64$; $p=.0006$). A single-sample t-test found that the new model was significantly more accurate than the old (M=0.46; $t(22)=2.42$; $p=.0245$). Staffing accuracy also improved at the service line level.

In Hospital A, the Medical/Surgical service line S-SbMAE metric improved from 0.51 to 0.45 while in the Critical Care service line F-SbMAE metric improved from 0.81 to 0.56. In Hospital B, the Medical/Surgical service line F-SbMAE metric improved from 0.86 to 0.50 while in the Critical Care service line F-SbMAE metric improved from 2.03 to 0.42. Staff was reallocated across each service line to provide coverage where needed as demonstrated in the S-SbMAE scores for the service line and the more balanced S-SbMAE values across the departments within the service line during the post-implementation period. The critical care service line experienced higher than normal volumes during the entire test period, which contributed to higher staffing challenges. The new model was able to accommodate the increased staffing pressures without a decline in staffing accuracy as the feedback loops worked as intended to resolve the imbalances. These scenarios provided opportunities to balance departments with extra staff with departments in need of staff as classic examples of engaging the balancing feedback loops of the subsystems. During the post-implementation period, there was an increase in staff floating across units, which was a topic often discussed during the project.

Table 29 – Staffing Results

Hospital	Service Line	Dept	Pre-State Average S-SbMAE	Post-State Average S-SbMAE	Diff	Binary 1=improved 0= not improved
A	Medical/Surgical	1	0.37	0.35	0.020	1
A	Medical/Surgical	2	0.35	0.46	-0.110	-1
A	Medical/Surgical	3	0.30	0.28	0.020	1
A	Medical/Surgical	4	0.60	0.48	0.120	1
A	Medical/Surgical	5	0.37	0.35	0.250	1
A	Critical Care	6	0.83	0.64	0.190	1
A	Critical Care	7	0.80	0.48	0.320	1
A	Emergency	8	1.05	0.65	0.400	1
B	Medical/Surgical	1	0.89	0.55	0.340	1
B	Medical/Surgical	2	0.21	0.37	-0.160	-1
B	Medical/Surgical	3	0.80	0.54	0.260	1
B	Medical/Surgical	4	0.93	0.51	0.420	1
B	Medical/Surgical	5	0.80	0.41	0.390	1
B	Medical/Surgical	6	0.73	0.59	0.140	1
B	Medical/Surgical	7	0.67	0.44	0.230	1
B	Medical/Surgical	8	0.67	0.24	0.430	1
B	Medical/Surgical	9	3.15	0.61	2.540	1
B	Medical/Surgical	10	0.60	0.66	-0.060	-1
B	Medical/Surgical	11	0.62	0.55	0.070	1
B	Medical/Surgical	12	0.00	0.00	0.000	0
B	Critical Care	13	0.55	0.17	0.380	1
B	Critical Care	14	4.56	0.88	3.680	1
B	Critical Care	15	0.98	0.23	0.750	1
B	Emergency	16	0.99	0.75	0.240	1

Additional benefits were captured via extensive nursing leadership interviews and meeting notes from the project. The most satisfying result from the majority of the leadership was being freed up from the office tasks of forecasting, scheduling and staffing providing more time to dedicate to more clinical based work. The nursing leaders felt they had more time to be on the floor rather than in their offices. Additionally, there was significant discussion about the preference honor rate achieved during the study period. Many nursing leaders recounted conversations with employees about schedule satisfaction and indicated this type of information would be useful to guide these conversations providing evidence of commitment to help employees achieve work-life balance.

Discussion

To the author's knowledge, this study represents the first attempt at a complete end to end assessment of the centralization of the labor forecasting, scheduling and staffing cycle in hospitals from a system perspective. The study accomplished two primary objectives. The first was the successful design and implementation of a centralized, systems based LPAC management model and the second was the demonstration of the model's positive impact on LPAC performance outcomes. The metrics used to assess phase performance demonstrated that improvement could be achieved from phase to phase of the cycle in the medical-surgical, critical care and emergency service lines with the centralization of functions and facilitation of knowledge sharing utilizing a systems approach. The new LPAC management model provided the necessary technical and social subsystem integrations for a successful transition where centralized role participants were able to perform their tasks with increased information and knowledge. The functions previously performed by nursing leaders at the department level to forecast, schedule, and staff departments were absorbed by the new roles in the central planning office freeing nursing leaders' time.

Observations demonstrated that subsystem communication increased as central resources facilitated more interaction between the service line and department leadership resulting in more proactive planning and a greater ability to adapt to changing circumstances. More options immediately available in the staffing phase to solve problems led to more flexibility. Central resources' view of the departments as sub-systems of larger systems provided better preparation to move staff to accommodate patient needs.

The social subsystem dealing with people, attitudes, and relationships is a more complex story. From a central planning office staff perspective, the artifacts identified and documented the social system integrations involving critical knowledge and informational elements for each of the three phases of the LPAC cycle. The developed model provided for the identification of these elements and guidance on where the elements can be found giving employees a knowledge map for completing their tasks. During the multiple schedule periods tested, the central planning staff was observed reaching out to the system and human sources of knowledge as they completed their work. We believe this led to a higher quality of input data to the individual processes resulting in higher quality outputs.

Social subsystem integrations from the perspective of the staff and nursing leaders, however, were not considered in this project. The centralization of labor-management functions can be a deeply emotional event for some staff and leaders. Advance and day-to-day work schedules are sensitive topics that can be directly dependent on the leader-staff relationship. We found during the project that these connections many times lead to staff favoritism as leaders tend to favor schedule requests from one employee over another. Employees that have more confidence to speak and negotiate or have better relationship building skills may get better schedules. We observed these scenarios in many circumstances where specific staff members were never working weekends or had “special” schedule preferences honored. The removal of nursing leadership from these decisions moved all staff to the point of equitable treatment. While this is typically interpreted to be a positive development ensuring equitable staff treatment, it can be a significant source of dissatisfaction for staff who lose their preferential treatment. Another challenge was some leaders’ feeling of loss of control. While the majority of leaders welcomed the relief from the back-office functions, there were some who had a difficult time relinquishing the control of the schedules and staffing. Leaders challenged with these feelings were difficult to move down the relationship development path with the central planning office. Often, these situations resulted in a lack of compliance. In these cases, department leadership ignored the central planning office and attempted to manage their staffing resulting in duplication of efforts. The result was poor staffing when duplicate staff members showed up for a given shift or duplicate staff members canceled. These instances weighed negatively on the post-implantation results. We identified that a compliance monitoring process is required for these scenarios to ensure that staffing direction is carried out successfully at the department level. A better understanding of how nursing leaders accept or do not accept this type of change will provide insight into new methods for implementing the culture change to result in less resistance.

The LPAC model artifact developed in this project is the core of our research. The model proved to be effective in giving back precious hours of time to nursing leaders through the removal of the back-office functions involved. Additionally, the model proved to be flexible enough to be adaptable to new circumstances and environments. An example is an ease with which the model was adjusted to account for intermittent department closures due to low volumes. We intended to create a model that allows for continuous improvement, adhering to one of the key components of HIM.

From the results provided, we conclude that there is an opportunity to achieve improved outcomes with a new model of LPAC management. While there is still significant work to more thoroughly test and investigate the full impact from a financial, staff/leader satisfaction, and patient outcome perspective, we believe that the improvements realized in this project bear out the contention that there is value in continuing this work. Workload forecasts can be improved leading to higher quality labor schedules that result in less last-minute staff adjusting that can result in avoidance of higher labor costs and staff dissatisfaction. While we expect in the long run, consistent labor schedules with a high rate of preferences honored leading to a better work-life balance will impact staff intentions to remain employed at the hospital; we believe more time is needed for the staff to accept and perceive these benefits outweighing other criteria that impact turnover. The time horizon of this project did not provide significant time to analyze these potential effects. We believe this research has provided the first insight into what hospital labor planning and allocation can look like at the system and subsystem level if centralized into well-defined roles with access to all relevant technical and social knowledge as input to the included tasks. It has also provided the first insight into how a systems approach can be used to better plan and allocate labor to meet the variable needs of patient flow through a hospital.

Limitations and Next Steps

There were several limitations to this study. Two hospitals consisting of a total of twenty-four departments across three service lines participated. The study needs to be extended to encompass Women's and Surgical services to cover the entire nursing environment. A second limitation is that the existing technology employed at the two test hospitals did not provide sufficient decision support capabilities or facilitation of technical and social integrations. RADs and IMCs were all created manually and were not available as part of role participants' day to day activities within a system. Based on the artifacts created in this project and recommendations provided by Harrison-Bronisnski, we envision an information system that accomplishes three primary objectives: role-based content management, automation, and collaboration/interaction management. Essentially role participants should see content specific to their role with the ability to switch between roles if they serve in more than one. This content needs to be editable and version controlled as role tasks many times involve the modification of data. The system should automate any mundane task that does not require human decision capability. Complimenting this recommendation is the need for decision support. The system must have decision support capabilities to facilitate decision making to leverage centralization economies of scale. Finally, the system should facilitate technical/social integrations to assist role participants with day to day activities including guidance on whom to interact with, what to exchange and activities to complete. Technology that supports the interdependencies of the subsystems and the facilitation of the technical and social integrations is a critical opportunity for supporting the new model to scale to a substantial size.

This project did not investigate the social integration impact from the perspective of the nurse leader and the staff and how these participants receive, process, and act within the new model. More work needs to be done to understand staff and nursing leader perceptions of the impact of the new LPAC management model on their day to day working lives may prove useful. The new model has the potential to increase staff movement (floating). Literature indicates that floating can have a negative connotation and may impact staff satisfaction. Since the systems approach ultimately leads to more staff movement across departments, a clear understanding of float perception would be helpful to improve the transition. One suggestion is a study focused on staff satisfaction with the LPAC model pre and post-implementation across multiple dimensions such as floating, preferences honored, communication, etc. Findings from this type of study could provide valuable information in the development of implementation strategies.

References

Ashforth, B. E., & Mael, F. (2016). Social Identity Theory and the Organization. *The Academy of Management Review*, *14*(1), 20–39. https://doi.org/10.1097/EDE.ObO13e31812e5535

Bertalanffy, L. von. (1950). An Outline of General Systems Theory. *The British Journal for the Philosophy of Science*, *1*(2), 134–165.

Boulding, K. (1956). General Systems Theory-The Skeleton of Science. *Management Science*, *2*(3), 197–208.

Castillo-Salazar, J. A., Landa-Silva, D., & Qu, R. (2016). Workforce scheduling and routing problems: literature survey and computational study. *Annals of Operations Research*, *239*(1), 39–67. https://doi.org/10.1007/s10479-014-1687-2

Chisholm, R., & Ziegenfuss, J. (1986). A Review of Applications of the Sociotechnical Systems Approach to Health Care Organizations. *The Journal of Applied Behavioral Science*, *22*(3), 315–327.

Defraeye, M., & Van Nieuwenhuyse, I. (2016). Staffing and scheduling under nonstationary demand for service: A literature review. *Omega (United Kingdom)*, *58*, 4–25. https://doi.org/10.1016/j.omega.2015.04.002

Glover, W., Li, Q., Naveh, E., & Gross, M. (2017). Improving Quality of Care Through Integration in a Hospital Setting: A Human Systems Integration Approach. *IEEE Transactions on Engineering Management*, *64*(3), 1–12. https://doi.org/10.1109/TEM.2017.2682267

Harrison-Broninski, K. (2005). *Human interactions: the heart and soul of business process management* (1st ed.). Meghan-Kiffer Press.

Hulshof, P. J. H., Kortbeek, N., Boucherie, R. J., Hans, E. W., & Bakker, P. J. M. (2012). Taxonomic classification of planning decisions in health care : a structured review of the state of the art in OR / MS. *Health Systems*, *1*(2), 129–175. https://doi.org/10.1057/hs.2012.18

Mark, B. A., Harless, D. W., McCue, M., & Xu, Y. (2004). A longitudinal examination of hospital registered nurse staffing and quality of care. *Health Services Research*, *39*(2), 279–300. https://doi.org/10.1111/j.1475-6773.2004.00228.x

Marsilio, M., Torbica, A., & Villa, S. (2017). Health care multidisciplinary teams: The sociotechnical approach for an integrated system-wide perspective. *Health Care Management Review*, *42*(4), 303–314. https://doi.org/10.1097/HMR.0000000000000115

Meyer, R. M., & O'Brien-Pallas, L. L. (2010). Nursing Services Delivery Theory: An open system approach. *Journal of Advanced Nursing*, *66*(12), 2828–2838. https://doi.org/10.1111/j.1365-2648.2010.05449.x

Mullarkey, M. T., & Hevner, A. P. (2015). Entering Action Design Research. In Donnellan B (Ed.), *DESRIST* (pp. 121–134). LNCS 9073.

Needleman, J., Buerhaus, P., Mattke, S., Stewart, M., & Zelevinsky, K. (2002). Nurse-Staffing Levels and the Quality of Care in Hospitals. *New England Journal of Medicine*, *346*(22), 1715–1722. https://doi.org/10.1056/NEJMsa012247

Ould, M. A. (1995). *Business processes : modelling and analysis for re-engineering and improvement*. Wiley.

Park, S. H., Blegen, M. A., Spetz, J., Chapman, S. A., & De Groot, H. A. (2015). Comparison of nurse staffing measurements in staffing-outcomes research. *Medical Care*, *53*(1), e1–e8. https://doi.org/10.1097/MLR.0b013e318277eb50

Pasmore, W., Shani, A., Francis, C., & Haldeman, J. (1982). Sociotechnical Systems: A North American Reflection on Empirical Studies of the Seventies. *Human Relations*, *32*(12), 1179–1204.

Phalp, K. T., Henderson, P., Walters, R. J., & Abeysinghe, G. (1998). RolEnact: role-based enactable models of business processes. *Information and Software Technology*, *40*(3), 123–133. https://doi.org/http://dx.doi.org/10.1016/S0950-5849(98)00047-0

Seo, W., Yoon, J., Lee, J., & Kim, K. (2011). A state-driven modeling approach to human interactions for knowledge intensive services. *Expert Systems with Applications*, *38*(3), 1917–1930. https://doi.org/10.1016/j.eswa.2010.07.124

Shani, A. B., Grant, R. M., Krishnan, R., & Thompson, E. (1992). Advanced manufacturing systems and organisational choice: sociotechnical system approach. *California Management Review*, *34*(4), 91–111. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=5615723&site=ehost-live

Silva-Martinez, J. (2016). Human systems integration: process to help minimize human errors, a systems engineering perspective for human space exploration missions. *Reach*, *2*(2–4), 8–23. https://doi.org/10.1016/j.reach.2016.11.003

Tarpey, R. J. (2017). Human Interaction Management Impact on Hospital Labor Planning. *Muma Business Review*, *1*(11), 125–139.

Tarpey, R. J. (2018). *Integrated Forecasting , Staffing , and Scheduling Outcome Measurement*.

Tarpey, R. J., & Nelson, M. F. (2009). Schedule Quality Assessment Metrics. *The Health Care Manager*, *28*(2), 145–158.

Tay, H. L. (2016). Examining Trade-Offs in Hospital Operations Using a Systems Approach in an Acute Care Setting. *International Journal of Business and Information*, *11*(2), 264–284.

Van Den Bergh, J., Belien, J., De Bruecker, P., Demeulemeester, E., & De Boeck, L. (2013). Personnel scheduling: A literature review. *European Journal of Operational Research*, *226*(3), 367–385. https://doi.org/10.1016/j.ejor.2012.11.029

Welton, J. M. (2011). Hospital nursing workforce costs, wages, occupational mix, and resource utilization. *Journal of Nursing Administration*, *41*(7/8), 309–314. https://doi.org/10.1097/NNA.0b013e3182250a2b

Chapter Seven: Overall Conclusion

The completion of this project resulted in the development of two main artifacts, the Systemic LPAC Management Model and LPAC Performance Metrics. The Systemic LPAC Management Model involved more than the centralization of tasks into a central service center. The detailed model documented each necessary role along with the technological and social components shown to be critical in the performance of the forecasting, scheduling and staffing phases of the LPAC. Role Activity Diagrams and Interaction Management Cards detailed each technological and social interaction involving critical data and information element flows necessary for the performance of the phase tasks. This documentation allowed for a thoughtful model design leveraging each value providing interaction.

The LPAC Performance Metrics provide for the accurate assessment of performance in each LPAC phase. Intermediate and end outcomes need to be monitored and assessed for role participants to take corrective or confirming actions. The dynamic nature of each of the phases requires continuous adjustment to support the final state of adequate department staffing per patient needs. Understanding status and performance allows for continuous improvement.

These artifacts were used to move the LPAC management model away from a sub-optimized department labor focus to a systemic focus seeking to balance labor across the numerous subsystems where needs exist. Managing labor at the system and sub-system levels better matched the patient flow management which occurred at these same levels. The new LPAC management model provided the opportunity for balancing feedback loops to operate more efficiently to keep the patient workload and staff capacity in balance. The overall goal of ensuring that the right amount of staff is in each department to accommodate patient needs adheres to the contention that adequate staffing impacts quality patient care. Departments in the post-implementation state operated with the correct staff to match their staffing plan on more occasions, then during the pre-implementation state. These findings were present on both sides of the equation. While departments operated less frequently short-staffed, they also operated less frequently over-staffed supporting labor cost control.

Each of the three phases of the LPAC across the two hospitals and three service lines experienced outcome improvement. It is important to note that while we believe significant improvement was achieved through the implementation of the new LPAC management model and operating from the systemic perspective, many variables can impact each LPAC phase. Internal factors such as staff vacancies, staff willingness to volunteer overtime, bed capacity, and patient throughput management are just a few that can impact how accurately staffed a department is for a given shift. Additionally, many external factors can impact results such as local incidents (e.g., large freeway accident) or regional incidents (e.g., flu outbreak, hurricane). One of the key objectives of the LPAC management model is to provide adaptability for processes to react to these less foreseeable and unpredictable events. The model is intended to provide for continuous model adaptations and evolutions.

The pre-implementation state of this project documented potential higher layer LPAC sub-optimization caused by attempted optimization at the lower layers as a result of siloed focus at the departmental level. The project essentially reversed these priorities and accepted the potential sub-optimization of the lower layers of the system in favor of optimization at the system higher layers. This concept was most visible within the scheduling function where less optimal schedule quality scores resulted from the optimization of schedule quality scores at the service line layer. One might argue that the system level approach, therefore, can readily lead to certain emergent behavior within the system defined as the undesirable or unexpected behavior of the higher layer system not found in the behavior of the individual lower layers.

We contend that emergent behavior is an inherent quality of any LPAC system due to the stochastic nature of the system itself. Past probabilities of patient movement combined with future probabilities determine workloads. Therefore, workloads exhibit a certain level of randomness leading to unpredictability. While managing the LPAC from the service line layer may introduce new behaviors not present at the departmental layer, the very nature of planning and allocating labor at the higher level provides a greater level of adaptability to overcome unexpected labor system behaviors. The concept of resource interchangeability at the service line layer supports this contention. The majority of skill sets at the service line layer can move between the departments within the service line as long as the movement is between like jobs requiring like skills and specialties. As an example, a general medical/surgical nurse can work on any medical/surgical floor assuming she has been oriented (trained) in each department from a logistics and process perspective. The same is generally true of a critical care nurse within critical care departments. This flexibility coupled with the centralized service level layer management provides mechanisms to adapt to potential emergent behaviors quickly. While this concept does not apply to highly specialized skills such as transplant or labor and delivery nurses, these patients are sufficiently isolated at the department level to reduce the requirement for this type of flexibility. This type of specialized workload to labor balance is achievable within the departmental layer.

Chapter Eight: Opportunities for Future Research

The project identified several opportunities for further investigation. The first area is technology. Existing technology does not fully support all of the concepts in higher layer LPAC systems management that we identified. Significant amounts of manual work were necessary to accomplish the many LPAC phase tasks at the higher layers. True scalability of the LPAC management model will be dependent on more advanced technology that can support the management of both technological and social interactions while providing decision support functions across each of the system layers.

The second area of opportunity is to more deeply investigate the sociological impact of the new LPAC model across the staff, leader, and patient perspectives. This project focused mainly on the social interactions necessary for the execution of the LPAC functions. An evaluation of the model change impact on employee satisfaction and patient experience was not feasible within the available time frame. An interesting additional study would look at the potential impact on different indications of employee engagement as well as potential impact to patient experience. While a more equitable approach to scheduling and staffing is assumed to be positive, there are indications that additional factors are important to consider such as the subsequent sociological impact of increased staff floating, less available premium overtime, and “forced” equality.

The third area of opportunity is to consider predictive analytics in the workload forecasting phase. This project did not investigate the performance of specific forecasting methodologies but rather left existing methodologies in place. These existing workload forecasting methodologies were historical extrapolation based and lacked consideration for cross sub-system patient movement. More advanced predictive analytics and business intelligence applied in this space could greatly improve the predictability of needed labor and further close the gap between the quality of the labor plan and the actual patient needs that need to be accommodated resulting in overall performance improvement.

The final area to consider is the potential impact of a Hawthorne Effect. The implementation of the new management model led to a large amount of observation and scrutiny within each nursing department. Future investigative work should test for the potential of the Hawthorne Effect by applying the same level of observation and scrutiny, but not changing the model and compare to a sequence of events that changes the model. The comparison of each group's results would provide valuable insight into the potential that better performance was attributed to the study itself rather than the model change. Equally interesting is the consideration that perhaps staff worked less hard with the increased observation and model change due to dissatisfaction in an attempt to discredit the change and revert to the prior way of doing business. Each of these different angles could provide valuable insight for future implementations.

Appendix 1 - Accepted Metric Calculation Formulas

Workload Forecasting Accuracy:

$$\boldsymbol{F - SbMAE} = \frac{\sum_{t=1}^{n} |A_t - F_t|}{n} \qquad t = 1, 2, \ldots n$$

Where:

- n = number of observations
- t = observation number
- A = actual labor bin
- F = forecast labor bin

Schedule Quality:

$$\boldsymbol{Completeness} = \frac{\sum_{i=1}^{n} |(S_i - R_i)|}{\sum_{i=1}^{n} (R_i)} \qquad i = 1, 2, \ldots n$$

Where:

- n = number of observations
- R = number of staff slots required
- S = number of staff slots scheduled

$$\boldsymbol{Commitment} = \frac{\sum_{y=1}^{P} \left(\sum_{x=1}^{E} f(M_{xy}) = \begin{cases} 1, H_{xy} < C_x \\ 0, H_{xy} \geq C_x \end{cases} \right)}{E * P} \qquad \begin{array}{l} x = 1,2,\ldots.\ P \\ y = 1,2,\ldots.E \end{array}$$

Where:

- x = each employee in the unit
- y = each week in a schedule period of P segments
- E = total number of employees in the department
- P = *n*umber of segments in schedule period
- M = total number of occurrences where the number of hours scheduled (H) is less than staff member's commitment (C)
- H = number of hours scheduled (per staff member)
- C = number of hours commitment (per staff member)

$$\textbf{Schedule Healtiness} = \frac{\sum_{y=1}^{E}(\sum_{x=1}^{N} U)}{E*N} \qquad x = 1,2,\ldots.\ N \qquad y = 1,2,\ldots.E$$

Where:

- x = each day in a schedule period of (N) days
- y = each employee in unit of (E) employees
- N = number of days in schedule period
- E = Number of employees in the unit
- U = occurrence of unhealthy shift patterns

$$\textbf{Preferences} = \frac{\sum_{y=1}^{E}(\sum_{x=1}^{N} R)}{E*N} \qquad x = 1,2,\ldots.\ N \qquad y = 1,2,\ldots.E$$

Where:

- x = each day in a schedule period of (N) days
- y = each employee in unit of (E) employees
- N = number of days in schedule period
- E = number of employees in the unit
- R = occurrence of employee request not honored

Staffing Accuracy:

$$\boldsymbol{S - SbMAE} = \frac{\sum_{t=1}^{n}|A_t - R_t|}{n} \qquad t = 1,\ 2,\ \ldots\ n$$

Where:

- n = number of observations
- t = observation number
- A = actual staffing bin achieved
- R = required staffing bin target

$$\boldsymbol{S - fMAE} = \frac{\sum_{t=1}^{n}|A_t - R_t|}{n} \qquad t = 1,\ 2,\ \ldots\ n$$

Where:

- n = number of observations
- t = observation number
- A = actual number of FTEs staffed
- R = required number of FTEs needed

Appendix 2 – Muma Business Review Reproduction Permission

Permission to reproduce use material from the following article published in the Muma Business Review is provided at the bottom of this page:

USF Muma Business Review
A publication of the Muma College of Business | University of South Florida

Volume 1, Number 11	Example Case Study	28 AUGUST 2017

Human Interaction Management Impact on Hospital Labor Planning

By
Richard Tarpey, University of South Florida

This study takes a distinctly unique approach to the hospital workforce planning (forecasting and scheduling) problem. The study is differentiated from precedent work in its focus on the structure of the work and the human interactions involved in labor planning, rather than strictly quantitative mathematical models and algorithms. Hospital labor planning involves many dimensions and levels of complexity. Within this complexity, we believe there are many improvement opportunities.

Labor cost is the single highest expense for hospitals. This case study describes utilizing Human Interaction Management to redesign work structure and process to improve labor forecasting and scheduling outcomes.

This study focused on examining human processes, interactions and work involved with forecasting workload and subsequent labor scheduling. The objective was to redesign necessary components to optimize human interactions, flow of information, and knowledge sharing in order to address the large amounts of complexity and variability.

The study concluded that a centralized role-process structure that facilitates and encourages more human interactions and feedback across the different roles resulted in more accurate labor forecasts, subsequently leading to more accurate labor schedules. We found that large amounts of critical knowledge and information was locked within the human participants who did not interact with other roles. There was a lack of a path for the critical information to flow across the roles where needed to successfully perform tasks. The drivers for the improvements were task focus and more information sharing leading to a richer collection of information and knowledge used as input to the work tasks. Redesigning work activities and roles resulted in better forecasting and scheduling outcomes as well as an additional benefit of freeing up clinical department leader time to focus on more patient and employee centric tasks within their departments.

Keywords: Human Interaction Management, Hospital Labor, Hospital Scheduling, Hospital Staffing, Workload Demand Forecasting.