MGH Internal Medicine Associates – Primary Care Redesign

by

Vaishal J. Patel
B.S. Mechanical Engineering, Cornell University, 2008

Submitted to the MIT Sloan School of Management and the Engineering Systems Division in Partial Fulfillment of the Requirements for the Degrees of

Master of Business Administration
and
Master of Science in Systems Engineering

In conjunction with the Leaders for Global Operations Program at the Massachusetts Institute of Technology

June 2015

Signature of Author Signature redacted
MIT Sloan School of Management and the Engineering Systems Division
May 8, 2015

Certified by Signature redacted
Retsef Levi, Thesis Supervisor
J. Spencer Standish (1945) Professor of Management, MIT Sloan School of Management

Certified by Signature redacted
David Simchi-Levi, Thesis Supervisor
Professor of Engineering Systems, Engineering Systems Division and the Department of Civil and Environmental Engineering

Accepted by Signature redacted
Munther A. Dahleh, William A. Coolidge Professor of Electrical Engineering and Computer Science Chair, ESD Education Committee

Accepted by Signature redacted
Maura Herson, Director of MBA Program
MIT Sloan School of Management

This page intentionally left blank.

MGH Internal Medicine Associates – Primary Care Redesign

by

Vaishal J. Patel

Submitted to the MIT Sloan School of Management and the Engineering Systems Division on May 8, 2015 in partial fulfillment of the requirements for the degrees of
Master of Business Administration
and
Master of Science in Systems Engineering.

Abstract

Internal Medicine Associates (IMA) is the largest primary care practice at Massachusetts General Hospital (MGH) with over 40 attending physicians, 60 residents, and 80 support staff that deliver care to more than 30,000 patients. The IMA is structured into seven pods that act independently to serve patients. Each pod consists of patient care providers and support staff that work collaboratively in a team structure. In particular, providers and medical assistants work closely together during the clinical session to meet patient needs. A lack of standardization in the practice's operations has contributed to inefficiencies that add to a sense of overload and burnout with the medical assistant staff. A detailed study of providers' clinical schedules revealed that individual clinical sessions are highly variable in terms of the number of concurrent clinical sessions per pod, session length, and number of patient appointments booked during this time. Providers in the IMA are part-time and create their clinical schedules based on personal preference and coordination with their other MGH related commitments. Variability in the schedule arises from many systematic, predictable, and unpredictable sources. Additionally, as part of a teaching hospital, IMA supports the educational training of over 60 Internal Medicine residents who hold a varying number of clinical sessions per week, depending on specific requirements of their residency program.

Coordinating and supporting provider presence consumes many resources, impacts medical assistant workload, and adds to variability within the practice. The project develops an optimization model to level-load the expected workload on medical assistants and other members of the medical care team by determining the clinic schedules of providers. The expected workload is measured by the number of concurrent sessions and expected number of patient visits per hour.

The project has developed an optimization model to suggest changes to the clinic schedule. Specifically in Pod 2/3, by strategically shifting 19.5% of provider sessions, we can achieve an 83% improvement in variability, as measured by the difference between maximum and minimum expected workload. Similar results are modeled for all pods in the IMA. The team has identified a pilot pod to test the model and is moving forward to operationalize the changes with IMA management.

Thesis Supervisor: Retsef Levi
J. Spencer Standish (1945) Professor of Management, MIT Sloan School of Management

Thesis Supervisor: David Simchi-Levi
Professor of Engineering Systems, Engineering Systems Division and
the Department of Civil and Environmental Engineering

This page intentionally left blank.

Acknowledgements

This work would not have been possible without the support and guidance of many people.

First, I would like to thank my academic advisors, Professor Retsef Levi and Professor David Simchi-Levi. Retsef's dedication to improving healthcare at MGH and personal guidance on this project have been critical to its success. His insistence on the highest standards and attention to detail pushed me to complete the project with a level of quality that would have otherwise not been possible. Thank you to David, for providing support throughout the project. Your insights into supply-chain flexibility and linear programming were invaluable.

The team at the MGH – MIT collaboration was instrumental in the success of this work. I would like to thank Dr. Peter Dunn for his deep understanding on the complexities of MGH and for directing my work to address the fundamental issues at IMA. To Bethany Daily, thank you for your support throughout my LGO project and fresh perspective you brought to the team.

I could not have completed this work without the support of the IMA. I would like to thank Dr. Sherry Haydock, Dr. Blair Fosburgh, Dr. Daniel Horn, Virginia Manzella, Sally Isles, Arlene Byrnes, and Diane Ford. Thank you for allowing me to be a part of the IMA and feel like a welcomed member of the team. Your dedication to improving patient care and devoting your time every week to share your insights, answer my questions, and support the MGH – MIT collaboration were vital.

Thank you to Cecilia Zenteno, whose limitless support throughout the project was no less than essential. I shudder to think where I would be without your technical expertise, attention to detail, patience, and leading the MGH happy hours.

Finally, I would like to thank my family, friends, and wife. To my family, thank you for your endless support throughout the MIT experience. To my friends, thank you for providing a sometimes needed distraction and allowing me to miss events as I completed this work. To Caslynn, I could not have joined MIT or completed this work without you.

This page intentionally left blank.

Table of Contents

List of Figures

List of Tables

List of Equations

This page intentionally left blank.

This page intentionally left blank.

1. Introduction

As the landscape of healthcare changes, Massachusetts General Hospital (MGH) must create new models of care. The primary aims of the MGH – MIT collaboration are to identify new models of evidence-based care and reallocate resources to best provide patient care. There are specific challenges in primary care as healthcare shifts from fee-for-service to population health management and medical home, focuses on cost savings, and integrates quality measures. Internal Medicine Associates (IMA) is the largest primary care practice at MGH and is facing similar challenges that are magnified by its size and complexity. The IMA has engaged with the MGH – MIT collaboration to be able to address the system redesign required to meet these emerging needs.

1.1. Background

1.1.1. Massachusetts General Hospital

Massachusetts General Hospital (MGH) is the third oldest hospital in the United States and the largest teaching hospital of Harvard Medical School. It is consistently ranked as one of the top hospitals in the world for its medical training and patient care. With an annual research budget of more than $750 million, MGH also conducts the largest hospital-based research. MGH and MGH Community Health Centers have approximately 48,000 inpatient and 1.5 million outpatient encounters annually, including 100,000 ER visits, 42,000 surgeries, and 3,600 deliveries.[1]

1.1.2. MGH – MIT Collaboration

Over the past nine years, MGH and MIT have collaborated to leverage analytical tools to design patient care processes across the hospital. MIT faculty, post-doctoral fellows, and LGO fellows have worked closely with MGH leadership to implement operational improvements in various departments throughout the hospital. For example, previous projects have addressed bed utilization in the perioperative environment, inventory management of surgical supplies, and patient flow through surgical units.

The first MGH – MIT project within Internal Medicine Associates focused on centralizing the prescription workflows into a centralized unit. The prescription process had previously been managed by all support staff as a part of many other daily activities they completed. The project proposed and implemented a centralized unit to fully manage the prescription refill process resulting in a more efficient allocation of staff resources and a more transparent workflow that is more robust to potential patient safety issues.

1.1.3. Primary Care

Primary care is the care provided by physicians specifically trained for and skilled in comprehensive first contact and care for patients with undiagnosed health concerns.[2] A primary care practice typically serves as the patient's first point of entry into the health care system. The primary care physician provides longitudinal care to patients that includes disease prevention, health maintenance, patient education, diagnosis and treatment of illness. Additionally, the primary care physician acts as the gateway to more specialized care depending patient needs and provides continuing care to the patient.

MGH Primary Cares services deliver care to more than 200,000 patients at 15 locations throughout the greater Boston area. MGH Primary Care facilities aim to provide, direct, and arrange medical services to meet all patient healthcare needs. MGH also owns and operates four community health centers whose mission is to provide services typically not found close to home and services to accommodate the needs of many different cultures.

MGH is Harvard Medical School's largest and oldest teaching hospital and as such, MGH Primary Care facilities provide comprehensive medical training to resident physicians. The MGH residency program has a world-renowned tradition of providing world-class training in academic medicine and biomedical sciences. Over a three to four year residency program, residents train through clinical rotations in the hospital's varied departments in both inpatient and outpatient settings.

1.1.4. Internal Medicine Associates

Internal Medicine Associates (IMA) is MGH's largest and oldest primary care practice with over 40 attending physicians, 60 resident physicians, and 80 support staff that deliver care to more than 30,000 patients annually. IMA is dedicated to improving healthcare through research, education, and innovation. In addition to primary care services, IMA offers on-site case managers, Blood/EKG Labs, vaccine clinic, and an internal Urgent Care unit.

The IMA is unique in that it is comprised of 7 independent "pods" that are each composed of healthcare providers and support staff. They work together in a team-based structure to deliver care to patients. There are between 5 – 9 providers and 4 – 7 support staff per pod. Providers, who provide direct patient care, are the attending physicians, nurse practitioners, and resident physicians. They work closely with support staff, who are the medical assistants and registered nurses, to meet patient needs. Notably, provider staff works on a part-time basis, while support staff works full-time. In 2014, 47 part-time attending physicians accounted for 17.6 full-time equivalent physicians.

Provider staff, who are responsible for a defined panel of patients, hold clinical sessions in their respective pods, where they see patients for their healthcare needs, on a part-time basis. Each provider manages the longitudinal health care of their specific panel of patients, ranging from 1000 – 2000 patients. Therefore, when a patient need arises, it is addressed by their specific primary care provider. On the other hand, the support staff works fulltime supporting those providers that are in session at any given time. The clinical session length, number of clinical sessions, and number of patients seen per provider is variable. Additionally, provider staff independently coordinate their weekly schedule based on their other responsibilities throughout the hospital and personal preferences.

Given the variability in provider presence and ultimately, the variability of workload on the pod over the week, there are times where there is excessive demand on support staff's capacity to support providers and meet the needs of patients. Due to daily workload variability, IMA support staff has experienced increasing dissatisfaction and burnout. Medical assistants, in particular, experience the effects of variability due to their close working relationship with supporting providers and working face-to-face with patients. There are times where medical assistants are overloaded with number of concurrent clinical sessions and number of patients to serve, which leads to overburdening a limited resource.

1.2. Project Overview

In the first phase of this project, we aim to understand resource use in the IMA through quantitative analysis. In order to learn and understand the workflows of the practice, comprehensive analysis of the existing data sources were complemented with shadowing sessions of the different members of the care team, as well as conducting individual interviews with all levels of staff. Existing electronic data sources, including MGH Primary Care Office Insight (PCOI) and IDX scheduling system were leveraged to analyze resource use. This effort provided insights into the complex nature of team-based care, resource capacity, and the adverse effects of under- and over-utilization of support staff. In particular, we focus on the workload of medical assistants due to their close working relationship with providers and patients. Medical assistants are an integral resource on the medical care team as they facilitate the flow of patients and support providers in delivering care. Additionally, the medical assistant's role is expected to become more critical in the future as healthcare shifts from fee-for-service to population health management.

In the second phase, we analyze the data gathered to identify the root cause of excessive and variable resource use in the IMA and to develop metrics to characterize utilization. Year-long, appointment level scheduling data along with observations of the physician clinic scheduling system resulted in identifying the focus area for operational improvements. We determine that variability in resource use is caused by variability in patient appointments per hour per provider, lack of standardized clinical session length, and clinical session schedule of providers. Provider staff members hold clinical sessions in their respective pods, where they see patients for their healthcare needs, on a part-time basis. On the other hand, the support staff works fulltime supporting those providers that are in session at any given time. Additionally, provider staff independently coordinate their weekly schedule based on their other responsibilities throughout the hospital and personal preferences.

In defining workload on support staff, we determined that it is measured most accurately by expected booked patient rates per hour per provider and number of concurrent provider sessions that they support. Booked patients are those patients that are expected to arrive for an appointment during a clinical session. Given the variability in provider presence and ultimately, the variability of workload on the pod over the week, there are times where there is excessive demand on support staff's capacity to support providers and meet the needs of patients. There are times where support staff are overloaded with number of concurrent clinical sessions and number of patients to serve, which leads to overburdening a limited resource.

The third phase develops an optimization model that schedules provider clinical sessions throughout the week for each pod in the IMA. The model's objective is to balance the expected workload for support staff while minimizing deviations from the current schedule. There are a number of systematic and practical constraints that are taken into account to ensure feasibility of the proposed schedule. For example, the model maintains existing working relationships between medical assistants and providers they work with to further facilitate team-based care. The optimization model will serve as a tool to analyze various scenarios and schedule providers in a more systematic way.

In the fourth phase, we analyze the proposed schedule against the current state in order to determine the tradeoffs between rescheduling providers and balancing expected workload. We analyze one pod in depth as a case study, as a representative example of the 7 total pods within the IMA. By shifting 19.5% of all provider clinics, Pod 2/3 can improve variability, as measured by the difference between maximum and minimum expected booked patient rate, by 83%. We discuss the magnitude of change required and the expected benefit to the pod. A pilot pod and accompanying guidelines will also be introduced in order to operationalize the proposed plan.

2. Literature Review

2.1. Trends in Primary Care

Primary care practice typically serves as the patient's first point of entry into the health care system and as the continuing focal point for all needed health care services. Primary care practices are organized to meet the needs of patients with undifferentiated problems that are cared for in the practice itself or referred to a subspecialty.[2] Primary care physicians provide longitudinal care to patients including disease prevention, health maintenance, patient education, diagnosis and treatment of illness. There is significant literature that suggests appropriate primary care can improve health outcomes and reduce costs.[3,4,5]

However, the rising costs and increasing complexity of health care options have highlighted the need for more efficient resources allocation. Improving productivity of already limited resources is vital to ensure the future of primary care today. Efforts to re-engineer workflows, leverage technology, and update operational policies to eliminate waste could yield widespread improvements in productivity.[6] Practices that do show high performance exhibit consistent characteristics: proactive planned care, shared clinical care among a team, streamlined workflows, and improved team functioning through co-location and team meetings.[7]

2.2. Team-Based Care

Historically, physicians operated mostly independently and were the primary care givers to their patients acting alone. However, the framework of health care has been continuously evolving from a model in which the physician is making house calls to team-based model. Today, the health care team is critical in delivering effective care to patients. Team-based care is defined as the provision of health services to patients by at least two providers who work collaboratively to accomplish shared goals to achieve high-quality care.[8] In this way, some tasks traditionally done by the physician is delegated to support staff.

There are a number of factors that influence the effectiveness of a health care team. First, it is important that there is clear distinction between who is on the team, and what their responsibilities are.[5] By having team member's roles defined and communication channels established, care teams can perform efficiently and with quality. Additionally, primary care teams that are more interconnected and less centralized and have a shared team vision are better positioned to deliver high-quality care.[9] Teams with higher cohesiveness are associated with better clinical outcome measures and higher patient satisfaction.

Studies indicate that involving all levels of staff in the health care team and empowering them to work to their full potential can increase satisfaction. This in turn can lead to increased productivity and decreased burnout and turnover.[10] Specifically, it is important to highlight the relationship between providers and medical assistants. Medical assistants have become the primary clinical assistants in family medicine offices, replacing registered and licensed practical nurses.[11] Although it varies by each practice, studies have found that medical assistants' proficiencies lie in patient flow and continuity and that their role is to facilitate the patient visit.[12] Medical assistants work closely with providers on a day-to-day basis to support their sessions with both clinical duties and administrative tasks. Trust and verification are key determinants in a successful working relationship between provider and medical assistant. A clinician's familiarity with the medical assistant as well as daily management helps develop that trust and verification.[13]

2.3. Patient Centered Medical Home

The Patient Centered Medical Home (PCMH) is a model of care that emphasizes care coordination and communication to transform health care delivery. The medical home encompasses a number of functions and attributes, such as comprehensive care, patient-centered, coordinated care, accessible services, quality, and safety.[14] PCMH is a model of primary care that is certified by the National Committee for Quality Assurance, NCQA. Primary care facilities that have achieved PCHM certification have shown to achieve better health care outcomes. Studies show that patients that identify with a PCMH have improved satisfaction, reduced errors, and improved quality of care.[15] MGH Primary Care, including IMA, aims to be certified as a PCMH.

2.4. Support Staff Capacity Planning

Literature discussed above highlights the importance and benefits of medical care team coordination on day-to-day operations and patient outcomes. A critical aspect to a successful medical care team is sizing the team with adequate resources to meet patient needs. From the support staff perspective, literature focuses on developing methods to set nurse staffing. A common technique is use a patient-to-nurse staff ratio to set the minimum number of nurses needed at a given time. In 1999, California passed legislation mandating patient-to-nurse ratios for its hospitals. The legislation was, in part, motivated by the perception that low nurse retention was related to burdensome workload and high levels of burnout.[16] A meta-study of existing literature found that low patient-to-nurse staff ratio is associated with lower failure-to-rescue rates, lower inpatient mortality, and shorter hospital stays.[17]

As suggested by existing literature, the primary criteria effecting nursing workload is the number of patients. There is little analogous work in the literature for medical assistants. Our work aims to characterize the workload and capacity of medical assistants in a primary care practice. We leverage similar patient-centric ideas currently applied to nurses in considering medical assistants.

2.5. Operational Improvements through Process Redesign

There is significant amount of literature on the application of operations management principles to the healthcare field. For example, lean manufacturing principles have been implemented in the Emergency Department (ED) to enhance the patient experience and staff satisfaction. Process redesign focused on patient flow analysis, empowering front frontline workers, and eliminating inefficiencies improved the number of patients seen and decreased length of stay within the ED.[18] Another study showed how process improvements that required little or no investment improved the turnaround time, from five to two days, for reports from the anatomical pathology lab.[19]

Operational improvements have been studied and implemented extensively within the various departments in hospital settings, and in particular, there has been focus on primary care practices. One study showed that strategic process improvements increased provider-patient continuity of care by 64% in one year.[20] Another study that focused on the prescription management process in a primary care setting showed that implementing lean tools increased the number of prescription renewals requests serviced per day.[21]

A 13-month lean implementation in the National Health Services, a publically funded care system in the United Kingdom, resulted in significant improvements in their primary care practices.[22] A value-stream map identified that 65% of 67 processes were waste and a 10-day time and motion study revealed that support staff performed on average 46% waste activities. Opportunities for waste reduction can be realized by implementing operational improvements such as standardizing processes, eliminating waste activities, and streamlining workflows. Another study conducted a cross-sectional survey of more than 60,000 patients who had recently visited a primary care practice in 34 countries.[23] The study found that comprehensive care in 23 of 34 countries offered a medium to high potential for improvement.

2.6. Optimization and Scheduling Tools

Discrete event simulation and linear optimization are common tools to analyze the current state of healthcare operations and suggest improvements. At MGH, simulation techniques are used to develop a detailed model to evaluate the impact of various operations changes.[24] Another study developed a simulation model to evaluate the downstream effects of scheduling rules and discharge process changes in the perioperative environment.[25]

Linear optimization is extensively used as a tool in scheduling and capacity planning decisions. The planning literature is rooted in aggregate planning studies relating to manufacturing. For example, a 1955 study developed a linear decision tool for production planning and sizing the workforce.[26] In healthcare, nurse scheduling is a common problem that is addressed through optimization. The nurse scheduling problem can be described as follows. The objective is to create weekly nurse schedules by assigning a number of possible shift patterns to each nurse. These schedules must satisfy a number of constraints including working contracts, individual preferences, and patient demand. There is significant literature that develop models to solve this problem. For example, one study creates an integer programming approach to schedule nurse shifts while another study uses a genetic algorithm.[27,28]

Our work applies a similar methodology to provider schedules, rather than nurses. Provider scheduling in the IMA is analogous to the classic nurse scheduling problem, in that, individual preferences, work contracts, and many other practical constraints must be incorporated in developing an optimal weekly schedule. We develop an integer program the optimally schedule provider clinical sessions.

This page intentionally left blank.

3. Current State and Challenges in the IMA

Internal Medicine Associates (IMA) is the largest primary care practice in Massachusetts General Hospital (MGH) in terms of patient visits and providers. It is organized around 7 independent "pods," where each pod consists of providers and support staff that work together to meet the needs of their patients. Pods are grouped into 3 teams:

- Team 1 consists of Pod 1ABC, Pod 5/6, and Pod 4C
- Team 2 consists of Pod 2/3 and Pod 4AB
- Team 3 consists of Pod 7/8 and Pod 9/10

Teams report to the administrative director, nurse director, and medical director. The administrative director and nurse director work together to manage support staff and day-to-day operations of the practice, develop and implement policies and procedures, support staff scheduling, and regulatory compliance. The medical director is responsible for the provider staff throughout the IMA, by overseeing quality of care, promoting population health management, and also maintaining regulatory compliance. Within each pod, there is an attending physician that is the medical unit chief and a nurse that is the nurse unit chief. These two roles also assist in managing day-to-day activities within their respective pods. Also, there is an urgent care pod within the IMA, IMA Urgent Care, where patients with urgent medical needs are seen. For the purposes of the analysis, IMA Urgent Care is not considered.

In the following discussion, we aim to characterize the pod structure and medical care team. Specifically, we discuss the complex interactions and workflows between the medical care team in order to successfully meet patient needs. Once we characterize how providers and support staff deliver care, we examine when providers see their patients. We explain the Master Grid Schedule as the baseline, standard schedule that providers commit to. However, there are many sources of variability that result in week over week deviations.

We develop a qualitative and quantitative understanding of the current performance of provider's schedules. The following analysis is based upon: (1) observations and interviews with the medical care team and (2) scheduling data. First, we characterize the interactions and workflows between various members of the care team as a snapshot in time. The IMA is a dynamic system in that many providers and support staff enter and leave the organization and for the purposes of the analysis presented, we consider the current state to be December 2014. Second, we use daily appointment level scheduling data to draw operational insights. All data was extracted from the IDX Scheduling System, which logs each booked appointment and records the scheduled date and start time, appointment type, patient information, booked provider, and appointment status (arrived, no show, canceled, etc).

3.1. IMA Pod Structure

Each pod is comprised of providers and support staff that work together to serve patients. There are between 5-9 providers and 4-7 support staff. Below we outline the roles and responsibilities, specific to the IMA, that comprise the team members of each pod:

Attending (MD): An attending physician is one who has completed their medical education and residency program to practice medicine. MDs have the legal and ethical responsibility to provide the primary source of medical care to their panel of patients.[29] The MD serves as the entry point for continual longitudinal care for patients and facilitates the need for specialty care.

Similar to other academic medical centers, attending physicians at MGH see patients part-time each week while managing their other hospital and research related activities. In February 2014, IMA had 47 attending physicians accounting for 17.6 full-time equivalent (FTE) physicians. A 1.0 FTE physician holds 8 clinical sessions and 2 administrative sessions. Therefore, physicians that are part-time and hold less than 8 clinical sessions are considered the appropriate fraction of 1.0 FTE. IMA attending physicians provide direct medical care for patients and give guidance and medical orders to support staff. The MD is the lead medical provider that coordinates with other health care team members.

Resident: A licensed doctor who has completed medical school, but is currently pursuing graduate medical training. As a part of resident's comprehensive Internal Medicine three-year training, they provide longitudinal care to patients in an outpatient ambulatory setting.[30]

In the IMA, residents have similar responsibilities to attending physicians, in that they are the primary providers to a panel of patients throughout their training. However in the case of residents, all medical care is supervised by their respective Preceptor. Residents rotate through many departments throughout the hospital during their residency as a part of their comprehensive training. They typically hold one clinical session per week called a continuity clinic. Also, during an Ambulatory Rotations, a resident holds one to two additional clinical sessions per week.

Preceptor: An attending physician that oversees the medical care provided to patients by resident physicians. The resident individually develops a medical care plan during the patient visit to address their needs. Prior to finalizing the plan, MGH attending physicians will provide consultation to the patient visit. First, the resident's assigned preceptor and the resident will consult in private to review the plan. Then, the preceptor and resident will visit the patient to finalize and complete the visit.

There is continuity in the relationship between the individual preceptor and resident physician in that they work together for a majority of that resident's clinical sessions. It is desirable that each resident-preceptor pair is scheduled at the same time every week to provide continuity. Each preceptor support the medical care of four residents. Although the resident is the primary care provider to their patient panel, the preceptor provides longitudinal care beyond the resident's three-year tenure.

Registered Nurse (RN): RNs are licensed professionals that typically have two or four year degrees that prepare them to perform many medical procedures such as administering vaccines and wound care.[31] RNs provide and coordinate patient care, educate patients about various health concerns, and provide emotional support to patients and families.

IMA RNs are the point of contact for patients to answer medication related questions between office visits and also proactively engages with patients who visited other specialties or the emergency room. RNs work closely with providers to check and manage blood pressure and diabetes between office visits.

Nurse Practitioner (NP): An NP has advanced clinical training beyond their initial professional registered nurse preparation. NPs have typically completed a master's or doctoral degree to prepare them for clinical competency to practice in primary care.[32]

Similar to residents and MDs in the IMA, NPs also provide longitudinal care to a panel of patients, but these patients are linked to an attending physician that oversees care. NPs serve patients that are less medically complex and may consult MDs for other issues. NPs are licensed to diagnose medical problems, prescribe medication, and perform yearly physicals. NPs also give medical directions to the support team.

Medical Assistant (MA): Medical assistants work alongside physicians, mainly in outpatient and ambulatory care settings.[33] MAs are non-licensed professionals that are cross-trained to perform both administrative and clinical duties.

In the IMA, MAs provide administrative support by greeting patients, updating patient medical records, and managing the controlled substance prescription process as directed by providers. MAs work closely with MDs, residents, and NPs by supporting their clinical duties, such as taking basic patient vitals, performing laboratory testing, and conducting other preventative health screenings.

Patient Service Coordinator (PSC): PSC is typically the first contact patients have with a primary care facility. A PSC is a non-licensed administrative support staff that schedules and coordinates patient appointments and ensures that all information needed by the medical care team is available.

PSCs and MAs work together to complete administrative tasks. PSCs check in patients and alerts the care team when they arrive for appointments. PSCs schedules further appointments and arranges for referrals when necessary.

3.2. Team-based Care Structure

Historically traditional approaches to healthcare were hierarchical in nature, in that a physician provided for all the needs of a patient and gave medical directions in a top-down fashion to support staff. In these models of care, the clinician operated in isolation and was responsible for all medical care related tasks. However, healthcare has evolved from the singular physician-centric care model to a team-based approach to care. A team-based approach leverages shared responsibility for patient care, clear communication channels, and maintains the ongoing relationship between the patient and care team.[34] In this way, all members of the care-team are more fully engaged in the pursuit of higher quality of care.

Team-based care is the cornerstone in how the IMA delivers care to patients, and it is essential to increase and develop it in the future. The roles described above in Section 3.1 coordinate with each other to meet the needs of patients. Figure 1 outlines the workflow of a typical appointment in the IMA:

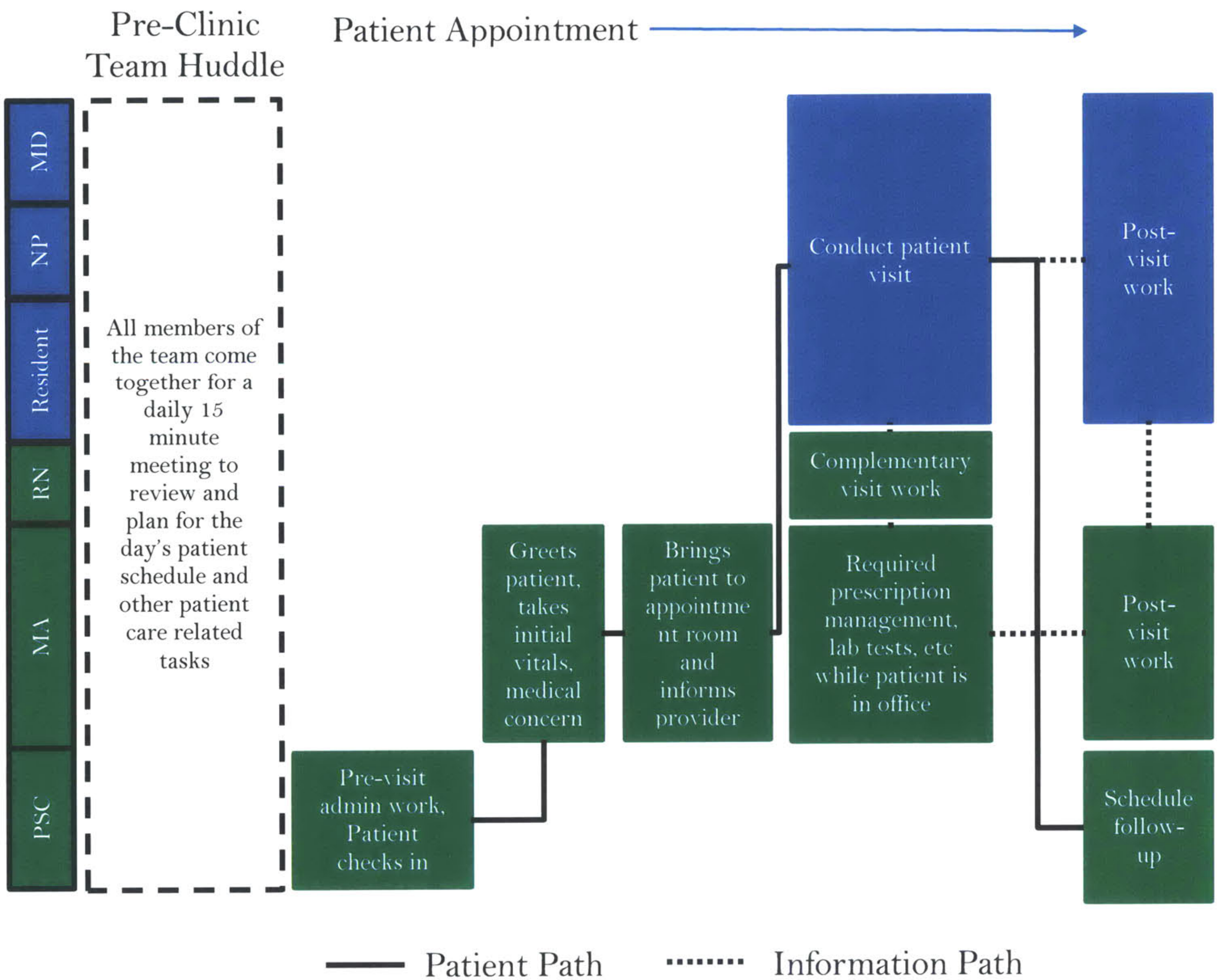


Figure 1 Appointment related tasks by members of the care team in the IMA.

A typical clinical session in the IMA begins with a "team huddle," where all members of the care team briefly meet to prepare and plan for that session's expected patient arrivals. The provider leads the discussion by detailing the expected needs of each patient appointment. The provider delegates and prioritizes clinical and administrative tasks to their respective support staff. These can span the full spectrum of medical needs, including preparing for vaccinations, laboratory tests, insurance paperwork, and specialty referrals. The huddle meeting serves as the starting point of the clinical session.

When the patient first arrives to their scheduled appointment, they are greeted and checked in by the PSC. The PSC logs into the IT scheduling system that the patient has arrived and if there are any urgent or emergency needs, the PSC will triage the patient accordingly. Once checked in, this signals to the MA that the patient is ready for the appointment to begin. The MA greets the patient and reviews medication and prescription related items, if applicable. The MA escorts the patient to an exam area where he or she takes vital signs, such as blood pressure, pulse, and weight. The MA, then, informs the provider the patient is ready and conveys any specific patient concerns or needs. The provider enters the exam room and addresses the medical needs of patient. An attending physician is assigned two exam rooms, while nurse practitioners and residents have one. Two exam rooms are used to minimize potential delays caused by the check-out process post-visit. Once the clinical work is completed by the attending physician, they can immediately attend to the next scheduled patient in the other exam room.

Outside of the exam room, the MA either completes tasks related to this appointment or will greet the next patient to begin the check in process. Once the provider has completed the patient's medical care, the MA and provider work together to complete any necessary post-visit tasks as needed. These can include administrative tasks, such as coding insurance forms or scheduling follow-up tests, and clinical tasks, such as laboratory testing or prescription renewal management. In this respect, medical assistants can have a variable workload depending on specific providers. Some providers delegate many tasks to the medical assistant with little oversight, while other physicians prefer to take proactive responsibility on administrative tasks. Then, the PSC finalizes the patient appointment and schedules a follow-up appointment if necessary.

3.3. Patient Panels

The medical team described in Section 3.2 works together to deliver care to the set of patients whose medical needs are managed by individual providers. Each provider is responsible for their specific panel of patients. A panel of patients are those patients that are linked to a specific provider, in that, the provider provides longitudinal primary care to them. In the following Figure 2, we note the size of provider panels normalized by their full-time equivalent (FTE). A physician that is 1.0 FTE will hold 8 clinical sessions and 2 administrative sessions. Therefore, we normalize individual provider's panel size by their FTE to characterize how many patients they care for. Figure 3 shows the distribution of panel size for residents by their academic year. We do not normalize by FTE for residents because all residents are expected to hold the same number of clinics per year (see Section 3.4.2).

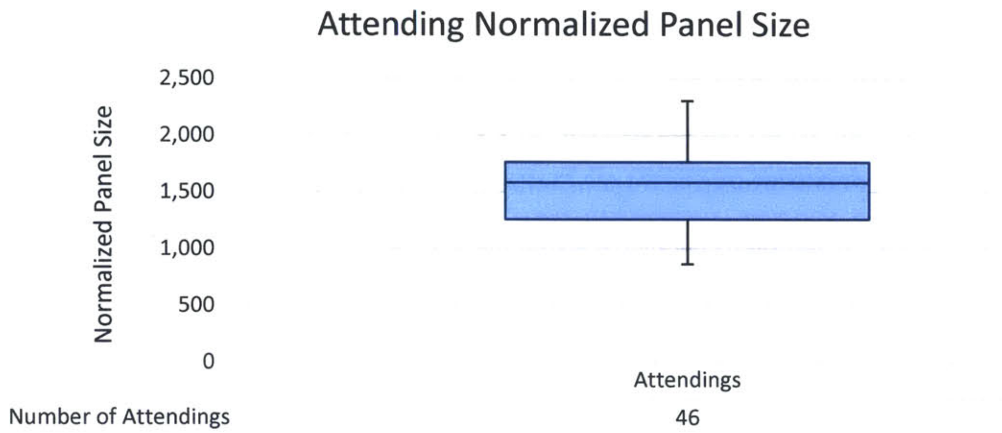


Figure 2 Distribution of attending panel size normalized by their FTE, based on February 2013.

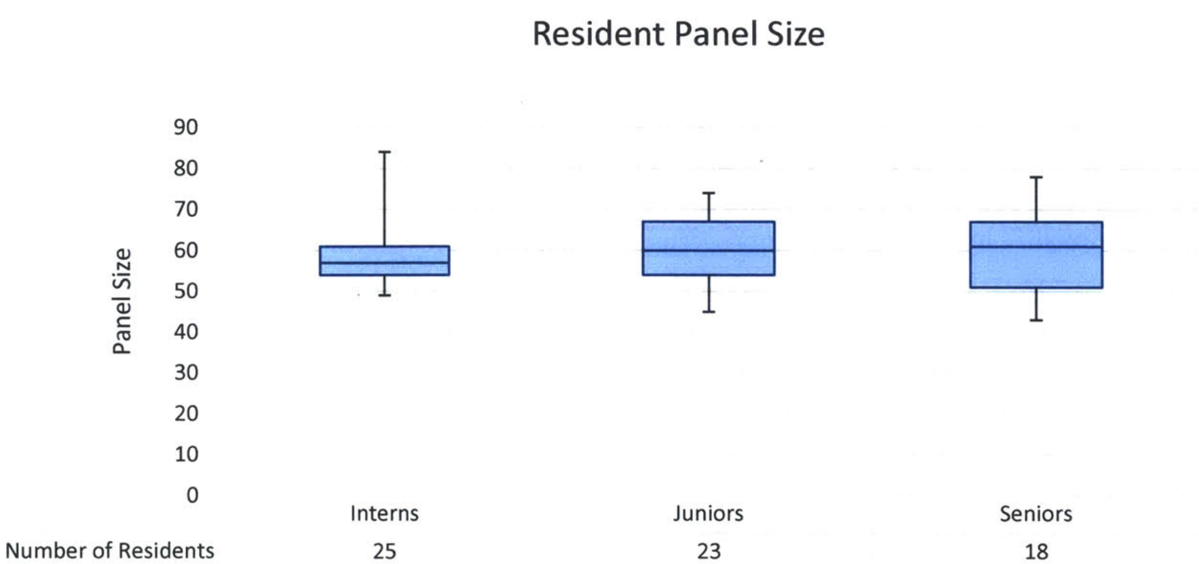


Figure 3 Resident panel size segmented by academic year, based on February 2013.

A number of factors account for the significant variation amongst providers in their normalized panel size. First, providers have varying levels of tenure at the IMA. There are some providers that have recently joined whose panel size may be smaller, while others have been practicing physicians for many years with many linked patients. Second, each provider has unique practicing preferences, where some providers typically schedule 15 minute appointments while others schedule 30 minute appointments. Additionally, there is variation in panel complexity between providers. Risk Adjusted Panel (RAP) score calculates a non-dimensional value that incorporates various dimensions affecting patient complexity such as, health insurance coverage, age, medical history, and frequency of healthcare visits. The calculation is outsourced to Ingenix, where each patient is assigned a RAP score annually based on their specific characteristics and subsequently, a provider's overall RAP score is the median score of their panel of patients. The RAP score is used to determine the relative complexity of provider's panels. The following Figure 4 outlines the distribution of RAP score of IMA providers and residents in 2013:

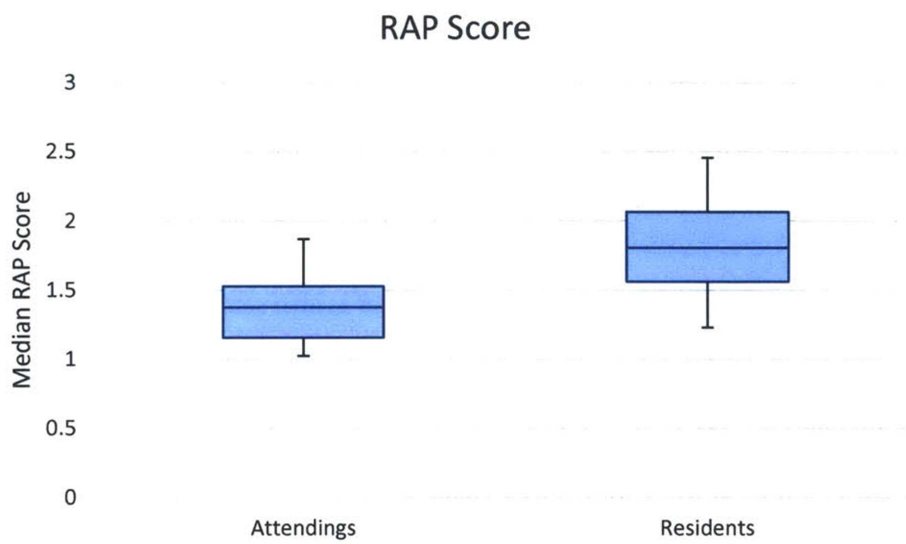


Figure 4 Distribution of provider panel RAP score by provider type, based on February 2013.

It is clear that there is a fundamental difference between resident and attending patient panels. Residents typically see patients in a primary care setting they have previously cared for in an in-patient setting elsewhere in the hospital. As a part of their education, residents rotate through various departments within the hospital and care for medically complex patients. When these patients require primary care services, they seek the resident that they have developed a relationship with and become linked to that resident's patient panel.

Patients with high RAP score typically require higher levels of medical care. Providers with a high RAP score patient panel will require higher levels of support from medical assistants, than providers with a less complex panel.

3.4. Provider's Schedules

The IMA is unique in that all clinicians are part-time, each of them holding between 1 – 6 clinical sessions per week, while support staff works full-time. In February 2014, 47 part-time attending physicians accounted for 17.6 full-time equivalent physicians. As discussed in Section 3.3, a provider that is a 1.0 full-time equivalent (FTE) is one that holds 8 clinical sessions and 2 administrative sessions per week (see Section 3.4.1 and 3.4.2 for session definitions). The weekly clinic schedule per attending physician and nurse practitioner is also self-determined. The provider staff coordinate their clinical sessions with their other responsibilities throughout the hospital and personal preferences. Conversely, residents have a pre-determined weekly schedule that coordinates with their other educational experiences.

In the following Section 3.4.1, we present the various session types and how the weekly clinical schedule is populated.

3.4.1. Master Grid Schedule

For planning and budgetary reasons, at the beginning of the fiscal year, IMA attending physicians and nurse practitioners commit to a number of clinical sessions per week they will hold and which session slot. This information is compiled into a template called the Master Grid Schedule. For each pod, there is a Master Grid Schedule that outlines the times during the week when each provider will hold a clinical session. Providers only hold clinical sessions within their pods. Resident schedules are governed by the MGH Internal Medicine Residency Program (IMRP) in conjunction with their other educational experiences.

The IMA is open for patient appointments between 9AM and 5PM every weekday and is closed for 11 holidays during the year. The following session types comprise the elements found on the Master Grid Schedule:

Clinical Sessions: Clinical sessions are held by attending physician and nurse practitioners in their respective pods to see and care for their patients. Patients are scheduled in appointments for 15 – 60 minutes to see their provider during these clinical sessions. There are a variety of appointment types, such as annual physical, follow-up, and urgent care. In the IMA, standard sessions are held for 3.5 hours either in the morning or afternoon. A standard "AM slot" is between 9:00AM and 12:30PM, and a "PM slot" is between 1:30 and 5:00PM. One hour between 12:30PM and 1:30PM is reserved for lunch, although it is frequently used as buffer time for appointments that are running over their schedule time.

However, the IMA is accommodating to individual provider preferences allowing for non-standard sessions lengths. Providers can work more or less than the standard 3.5 hour session and have variable start and end times. Some providers also hold their clinical session through the lunch hour, either by design or when the session is running past schedule. Variability in individual provider preferences in clinic start times and patient appointments scheduled during the lunch hour puts a burden on support staff resources.

Administrative Session: Providers hold administrative sessions in the IMA where they use an exam room as an office to complete administrative tasks. These sessions do not require any assistance or attention from support staff by definition. However, providers frequently allow patients to be scheduled during their administrative sessions for emergent reasons which results in an unplanned demand for support services.

3.4.2. Residency Program

The residency program is a hallmark of MGH as an academic medical center. As a part of the three year residency program, resident physicians rotate through various departments. Residents also provide longitudinal care for a panel of patients by seeing patients during their resident continuity session which is scheduled onto the Master Grid Schedule.

Resident Continuity Session: In order to provide timely access to care for patients, residents hold a weekly continuity session in their respective pods. This is a clinical session, similar to attending physicians and nurse practitioners, where they provide care to their patients. The continuity session is scheduled into the Master Grid Schedule at the beginning of the academic year and is expected to be completed every week. The continuity session is a full 3.5 clinical session and residents book either 30 or 60 minute appointments.

For one third of the academic year, residents are on an Ambulatory Care Rotation (ACR), where residents hold an additional two clinical sessions per week. This provides increased access to care for their patients and enriches the resident's educational experience. The two additional clinical sessions are scheduled by an outside vendor, Stottler Henke's proprietary software Aurora.

Stottler Henke coordinates all MGH IMRP educational experiences for each resident. When residents are not on ACR, they are in in-patient rotations which require a "call schedule." A call schedule is where one resident must be available for 24 hours to deliver medical care to the in-patients of the hospital. For the day prior and day after a resident is on call, they cannot be scheduled for a clinical session in the IMA due to work hour restrictions. The Aurora scheduling system manages the complexities of the resident schedule for both in-patient and out-patient rotations.

Although residents are responsible for their panel of patients, ultimately, the care residents provide is governed by a specific attending physician, called a preceptor (see Section 3.4.2). The preceptor session is also scheduled onto the Master Grid Schedule.

Preceptor Session: Attending physicians that are Preceptors in the IMA hold preceptor sessions. These are sessions where a preceptor is scheduled to an exam room or office to be available to resident physicians for consultation on their respective patient appointments. In a private room, preceptors and residents discuss the patient's needs and the proposed medical care. After agreeing on the course of medical care, preceptors and resident physicians will consult with the patient together, in the patient exam room, to further reinforce the resident's clinical education. Once complete, the preceptor will return to their office to be available for another resident. Due to the additional mentoring needed, resident appointments are typically twice as long as attending appointments where all appointments are either 30 or 60 minutes.

3.4.3. Medical Assistants

It is important to note that Figure 1 illustrates the average patient appointment workflow and many modifications are made based on unique patient needs and provider's individual practices. The team works closely and communicates often to adapt to changing needs throughout patient appointments. In particular, the medical assistant is typically the provider's primary resource for support related tasks and therefore, the relationship between the provider and MA is a critical link in a successful medical care team. As noted in Figure 1, there are important face-to-face information flows both in-visit and post-visit between provider and MA.

For each particular clinical session a provider holds, they will often be supported by the same MA. Over the years, strong working relationships form between individual providers and MAs that consistently work together. Continuity within the care team ensures that the team functions seamlessly and efficiently and in this way, medical assistants are a critical resource in delivering positive patient outcomes. Each pod aims to create a consistent care team for each provider for each of their clinical sessions.

Due to the part-time nature of providers, MAs frequently support many providers at one time, although overall, they look to follow a 1:1 model between MA and provider. Given the level of responsibilities MAs have with patient-facing, provider-facing, and other related work, individual MAs can support a maximum of two concurrent clinical sessions. Medical assistant's responsibilities are also expanding as the future of healthcare shifts to team-based care. Medical assistants taking on a greater role in the medical care team is a critical link as the IMA aims for PCMH certification. Additionally, an MA's daily workload is correlated with the expected number of patients. Limiting the number of concurrent sessions supported by each MA serves as a proxy to limit the number of patient appointments per hour per MA.

3.4.4. Sources of Schedule Variability

Section 3.4.1 and 3.4.2 describe the elements comprising the Master Grid Schedule. With no variability in provider behavior, each pod in the IMA can expect 100% adherence to this schedule week over week. However, there are systematic, predictable, and unpredictable sources of variability that create deviations. In this section, we examine the qualitative sources of variability.

Systematic Variability

There are a number of sources creating variability in the schedule inherent to the system design of the Master Grid Schedule. First, attending physicians and nurse practitioners determine their clinical schedule soley based on their individual preferences. Each provider can have other hospital responsibilities that include specialty work, research interests, and teaching commitments that they must balance with clinical sessions in the IMA. This leads to little oversight over the aggregate Master Grid Schedule for each pod. This results in an unbalanced workload over the week where there are session slots scheduled to have more concurrent clinical sessions than others.

Clinical session lengths are also a source of systematic variability. A standard clinical session is 3.5 hours with prescribed start and end times, but many providers modify their schedules. A number of providers conduct a non-standard session length and schedule patients into the 1 hour lunch break by systematic design. This leads to increased dissatisfaction from staff that support these provider sessions.

When clinical sessions need to be rescheduled for the predictable and unpredictable sources identified below, these are scheduled into unplanned session slots. There is little systematic or management oversight in when make-up sessions are scheduled. Unplanned clinical sessions lead to excess demand for support staff that adds to the sense of burnout.

Resident continuity sessions are always scheduled for the PM slot to avoid conflicts with other educational requirements and work hour restrictions, which results in an unbalanced load over the day. During a resident's ACR, the two additional clinical sessions cannot be scheduled on the AM slot on Thursdays, but can be scheduled for the AM or PM slot on other days of the week. Residents have a mandatory didactic session outside the IMA as part of MGH IMRP requirements during this time.

Predictable Variability

Predictable schedule variability results from sources that are unavoidable, but can be predicted in advance. For example, IMA allows attending physicians and nurse practitioners 8 weeks of vacation that can be used any time, resulting in 44 required weeks of clinical time. Vacation times are typically planned in advance and clinical sessions that fall during this time are rescheduled proactively. Also, there are many yearly conferences and holidays that affect provider schedules. Providers attend conferences depending on their research or educational interests that are known in advance.

Predictable schedule variability results in many clinical sessions that need to be rescheduled in order for providers to meet their clinical session requirement. These sessions are rescheduled on an as needed basis based on provider preferences, which further exacerbates the systematic variability.

Unpredictable Variability

Unpredictable schedule variability results from unavoidable and unpredictable sources. The primary drivers are sick days, last-minute schedule changes, and changes to resident schedules. Unpredictable variability is inherent to the working environment and cannot be planned for, where clinical sessions are rescheduled retroactively. Given many educational requirements for residents, unpredictable changes to their clinic schedule occurs often. The outside vendor, Stottler Henke, manages the in-patient and out-patient resident schedules, but can create errors by scheduling clinical sessions around the call schedule. The work hour restrictions are closely monitored by IMA management and manually changed when necessary.

A significant source of unpredictability stems from last-minute appointment cancellations and no-shows. A no-show appointment occurs when a patient does not arrive to a previously scheduled appointment with no notice. No-shows and last-minute cancellations disrupt the continuity of the session and further contributes to the stress of the medical team that had prepared for the patient visit.

3.5. Data Analysis of Master Grid Schedule

In this section, we present a data analysis on the performance of the schedule from June 2012 - June 2013. The Master Grid Schedule provides the standard, baseline schedule that all providers should adhere to in the absence of variability. We examine how the sources of variability discussed in Section 3.4.4 result in week over week deviations at the pod level.

As discussed in Section 3.2 and 3.4.3, MA workload is determined by the number of concurrent clinical sessions that they support and the number of booked patient appointments. A primary driver for increased dissatisfaction and burnout among the support staff is variability in provider schedules resulting in more than two concurrent clinical sessions. We examine the occurrences schedule variability resulting in overburdened MA support.

3.5.1. Adherence to Master Grid Schedule

At the beginning of every fiscal year, IMA attending physicians and nurse practitioners commit to their clinical schedule. In Figure 5 below, we examine the adherence to scheduled clinical sessions. For all clinical sessions that a provider commits onto the Master Grid Schedule, it portrays how often they adhere to the 44 clinical session requirement in the year. Each data point in each boxplot represents the adherence percentage for each providers' clinical sessions where all sessions are treated individually.

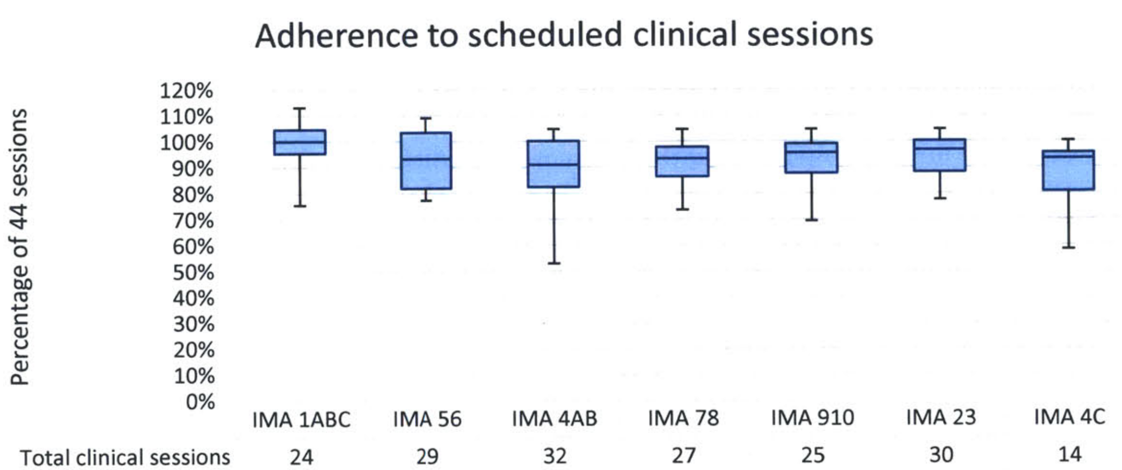


Figure 5 Adherence to scheduled clinical sessions by all providers in each pod.

Appointment level data over June 2012 – June 2013 is used to determine whether a provider was in session or not. If a provider had a scheduled appointment where there was a patient arrival or patient no-show, then the provider was considered to be in session. We use this method as a benchmark for other analysis throughout this section. Every occurrence of a clinical session was tabulated for each provider for each session slot and compared to the 44 session standard as a percentage. It is possible for greater than 100% adherence if the provider did not take 8 weeks of allowed vacation time. The distribution of adherence to all clinical sessions per provider is shown in Figure 5.

As an example, a provider in IMA 1ABC is scheduled to be in session for the Monday AM slot per the Master Grid Schedule and during June 2012 – June 2013. In reality, the provider held 33 clinical sessions during that time resulting in 75% adherence.

The majority of providers in each pod adhered to their clinical session between 90-95%, that is, 2-4 clinics were not held during their prescribed time slot. However, there is significant variability in provider behavior. Pod 4AB and 4C exhibit significant variability in provider presence during their assigned session times. As a result, providers must reschedule their sessions outside of their standard schedule to meet the IMA standard. We distinguish between two types of sessions in which providers see patients outside their scheduled clinical time, make-up sessions and extra sessions.

3.5.2. Make-Up Sessions

For providers that do not meet the 44 clinic sessions requirement, they usually schedule make-up sessions at their preference assuming there is exam room capacity. The following Figure 6 illustrates the distribution of clinical sessions that were scheduled outside of the provider's standard Master Grid Schedule.

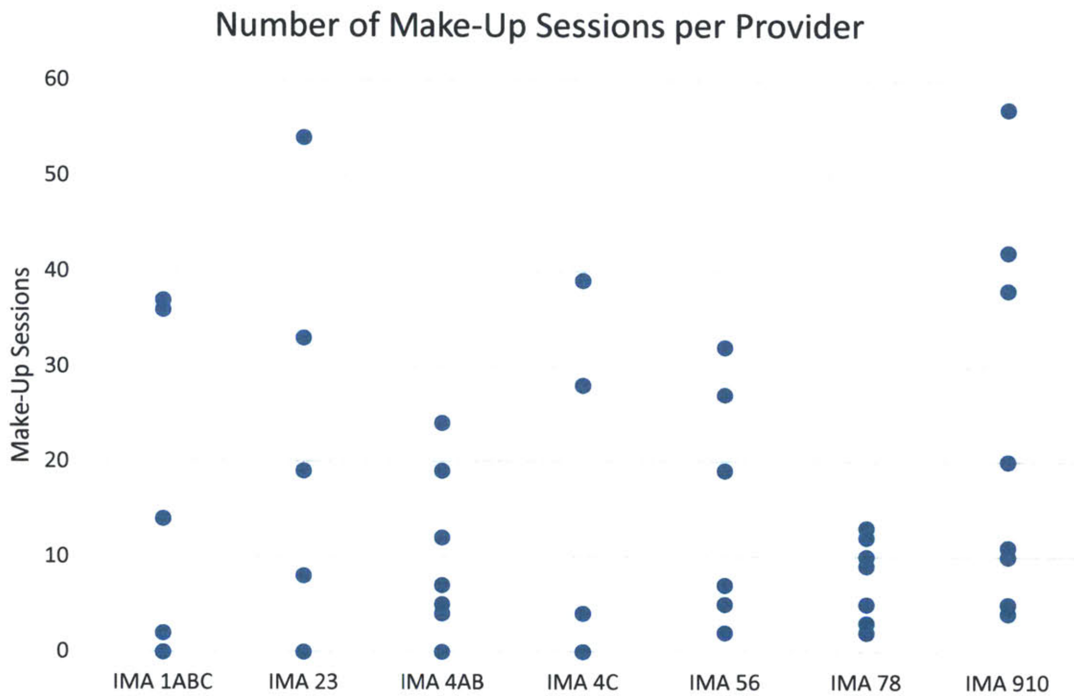


Figure 6 Distribution of make-up sessions per provider.

Similar to the analysis of the provider's adherence to the Master Grid Schedule, appointment level data is used to determine which session slot a provider is in session. Every occurrence of a clinical session outside of the standard Master Grid Schedule is counted. It is important to note that providers frequently schedule last-minute patients for emergency reasons at predictable times outside of their standard schedule. These predictable times occur regularly enough that we consider them as a special case. The particular occurrences were separated from this analysis (see Section 3.5.3). Figure 6 shows the distribution of the number of make-up sessions per provider from June 2012 – June 2013. Each point represents the total number of make-up sessions scheduled by a provider.

Many providers hold 10 – 30 make-up sessions in total throughout the year, with high variance in individual provider behavior. Figure 6 outlines the sum of all make-up sessions held by a provider, these do not necessarily occur during the same session slot. The providers in Pod 4AB and 7/8, for example, adhere to their clinical schedule more closely on a week to week basis than those providers in Pod 1ABC and 9/10. The excess number of make-up sessions is an unpredictable demand on support staff resources. Given the widespread occurrence of make-up sessions among providers, support staff in each pod is subject to the resulting unpredictable workload. Considering both provider adherence to the Master Grid Schedule and make-up sessions, many providers exceed the 44 session standard. Increasing patient accessibility and not fully utilizing vacation time are the primary drivers of providers holding more than 44 sessions.

3.5.3. Extra Sessions

Extra sessions are the occurrences of scheduled patients outside of the provider's standard clinic time that occurred more than 20 times in a given session slot during the year. This captures the semi-regular demand on support staff that is not accounted for on the Master Grid Schedule and are in addition to make-up sessions discussed in Section 3.5.2. Emergent patient needs are frequently met by leveraging the provider's presence in the IMA through their administrative time. Their standard clinical sessions can be fully booked by other patients many weeks in advance and in order to increase patient accessibility for urgent needs, providers will schedule appointments during their administrative time. Also, some providers habitually schedule patients outside both their clinic and administrative time when there is excess exam room capacity to see these urgent patients. The following Figure 7 outlines the number of extra sessions held per provider and Figure 8 outlines the number of patients scheduled during this time.

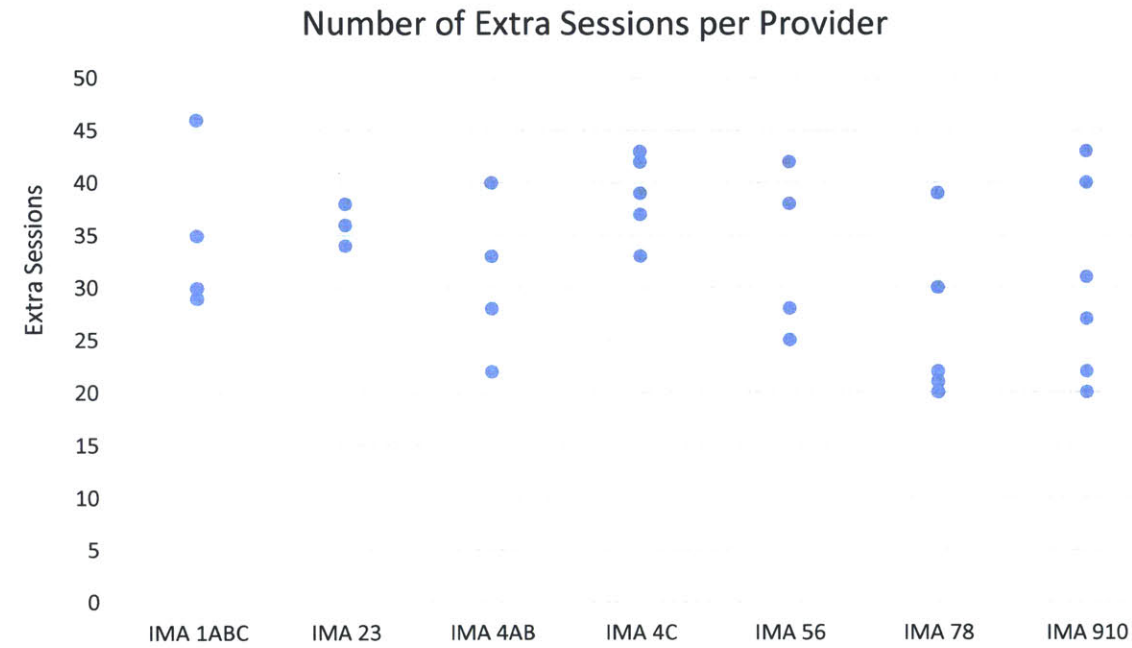
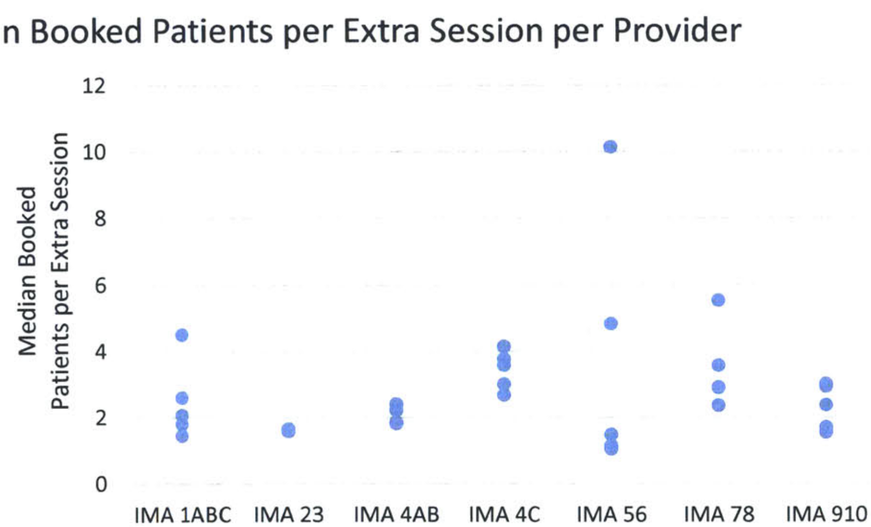


Figure 7 Number of additional sessions per provider that are considered "extra sessions."



	IMA 1ABC	IMA 23	IMA 4AB	IMA 4C	IMA 56	IMA 78	IMA 910
Number of Extra Sessions	5	3	4	5	5	5	7
Number of Providers	2	1	2	2	3	2	5

Figure 8 Distribution of the number of booked patients per extra session.

Appointment level data is used to determine when a provider is holding their clinical session. Every occurrence of a patient arrival that occurred outside their standard clinic time is tabulated. Then, we aim to identify those session slots are semi-consistently used, defined as greater than 20 occurrences, by each provider. Figure 7 illustrates the distribution of extra sessions by providers that use extra sessions per pod. It is important to note that those providers that do not hold extra sessions are not included in the distribution.

For each identified occurrence of an extra session, we tabulate the number of patients scheduled for that time. Figure 8 illustrates the distribution of total number of patients scheduled during a provider's extra session.

There is significant variability between providers in how many extra sessions they hold. Some providers use their administrative time extensively to see patients, while others rarely schedule patients during this time. It is important to note that extra sessions are not necessarily standard 3.5 hour clinical sessions and typically, are shorter. Figure 8 shows that typically 1 - 3 patients are scheduled during an extra session. Extra sessions are an unplanned workload on support staff that must be accommodated when these sessions occur. Extra sessions are used semi-regularly by some providers, but are not captured on the Master Grid Schedule and subsequently, are not incorporated in allocating support staff resources.

3.5.4. Resident Schedules

The Accreditation Council for Graduate Medical Education, ACGME, is an international accreditation organization that accredits residency programs to improve healthcare and advance resident education.[35] ACGME requires that residents conduct 130 outpatient ambulatory care clinic sessions during their residency, which translates to approximately 43 clinical sessions per year.[30] In order to meet this requirement, residents are scheduled a continuity session once per week and Ambulatory Care Rotations (ACR) for additional clinical sessions at various times.

Resident Continuity Session

At the beginning of each academic year, residents are assigned to a weekly clinical session called the continuity session. In order to increase patient accessibility, each resident is scheduled for a clinical session during this session slot per week for the entire academic year in order to meet their minimum requirements. However, many predictable and unpredictable sources of variability result in cancelling or rescheduling the continuity clinic. For example, instances occur where a resident is scheduled to be on 24-hour call in the in-patient department and is consecutively scheduled for a continuity clinic. In this predictable situation, the continuity will be cancelled and possibly rescheduled due to work hour restrictions. Also, unpredictable sources arise, such as provider illness or last-minute vacation requiring cancellation. The following Figure 9 shows the adherence percentage of each resident to their respective continuity clinic session slot.

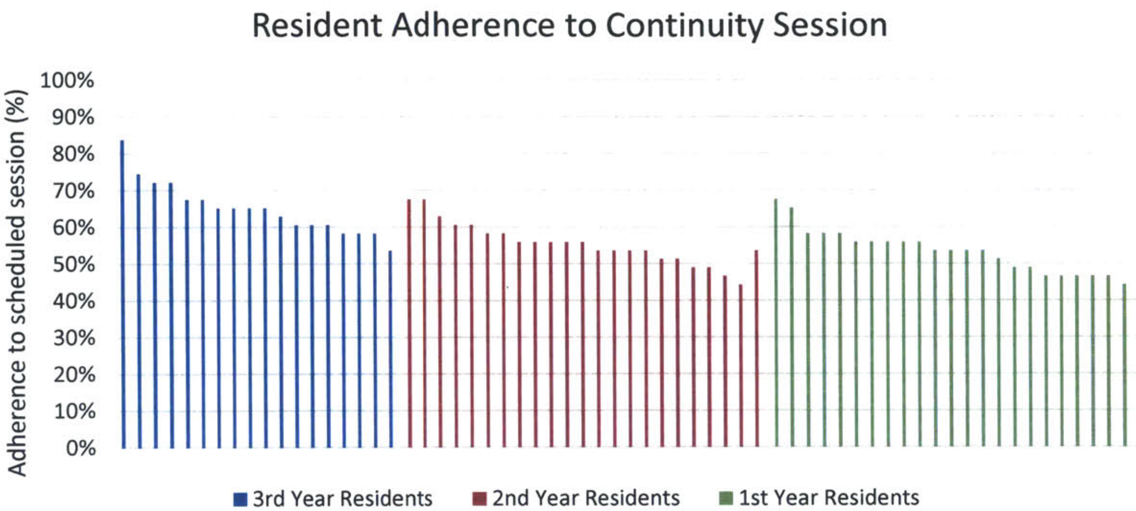


Figure 9 Adherence to assigned continuity session per resident.

Each data point represents the percentage of clinics held by each resident on their assigned continuity clinic time against the 43 session standard. There is significant variability between individual residents and overall, all residents have less than 90% adherence to their scheduled time. In order to meet the minimum 43 session requirement, the remaining clinics are manually rescheduled on an as needed basis. Available exam room capacity, conflicts with other educational requirements, and preceptor availability are the primary considerations when rescheduling a resident session and support staff capacity, however, is typically not a factor. Given that the majority of residents require schedule changes to 40 – 60% of their continuity sessions, this puts an uncontrolled and unplanned demand on the support staff.

Ambulatory Care Rotations

In addition to continuity sessions, MGH residents rotate through 6 ACR blocks during the academic year. In each ACR block, four additional clinical sessions are scheduled, outside of their continuity session during a block, a two week period. In sum, a resident will hold 6 clinical sessions during an ACR block. Each resident has a unique block schedule, in that, individual residents are on different block schedules at any given time during the academic year. The following Figure 10 outlines the distribution of the number of residents on ACR per pod per week.

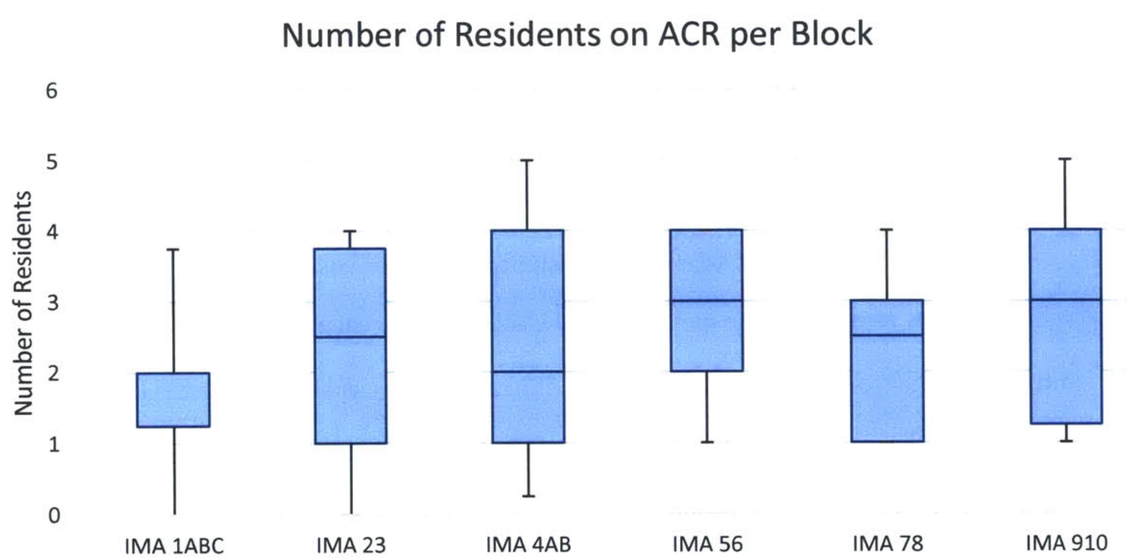


Figure 10 Number of residents that are on ACR per block per pod.

There is significant variability in the number of residents on ACR, which directly translates to variability in the number of sessions medical assistants must support during the week. For example, during a peak block in IMA 910, there were a total of 30 resident clinics, while the minimum was 6. Scheduling the additional clinical sessions are viewed similarly to rescheduling the continuity clinic, in that, preceptor availability, exam room capacity, and educational requirements are the primary considerations. The ACR sessions can introduce significant variation into the number of clinical sessions that require the support of medical assistants.

3.5.5. Appointment Status

As noted in Section 3.4.4, a primary source of unpredictable variability is last-minute cancellations and no-show appointments. In the following Figure 11, we outline the no-show rate by provider type from historical schedule data from June 2012 – June 2013. The no-show rate is defined by Equation 1:

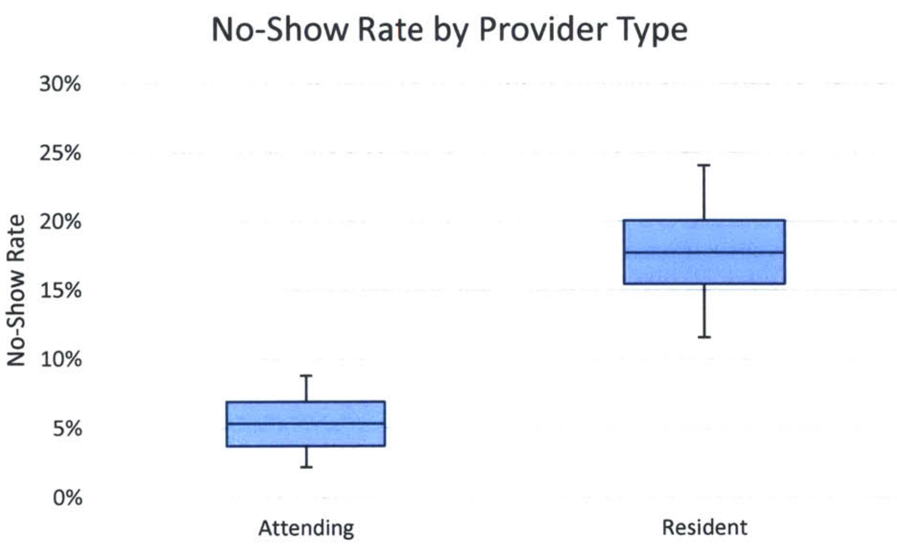


Figure 11 No-show rate of providers from June 2012 - June 2013.

$$No - Show\ Rate = \frac{Number\ of\ No - Show\ Appointments}{Number\ of\ Arrived + No - Show\ Appointments} \quad (1)$$

Resident patients are more likely to be medically complex, shown in Figure 4, and exhibit a higher no-show relative to attending patients. No-show appointments disrupt the medical care team and adds to the support staff workload. Support staff take the lead responsibility in investigating a no-show appointment. This leads to unpredictable demand on support staff because effort is spent making phone calls, filing paperwork, and rescheduling appointments.

3.5.6. Medical Assistant Capacity

To ensure adequate resource utilization and minimize patient wait times, IMA management has dictated that a MA should not support more than one attending, two nurse practitioner, or three concurrent resident clinical sessions. The difference in standard is due to attending physicians that typically schedule more patients per hour compared to nurse practitioners and residents. Additionally, residents require less support from medical assistants because they check-in and take vitals of their own patients. Given the level of responsibility MAs have with patient-facing, provider-facing, and other related work, individual MAs can only support the prescribed number of sessions depending on the provider type before their capacity is overloaded. It is important to note that medical assistant capacity is a function of (1) the expected booked patient rate and (2) the number of concurrent sessions.

The clinical session variability discussed above results in significant day over day and week over week schedule differences that negatively affects medical assistant capacity standards. In the following Figure 12, we show the distribution of the number of medical assistants required per session given that session's provider presence over the year June 2012 – June 2013.

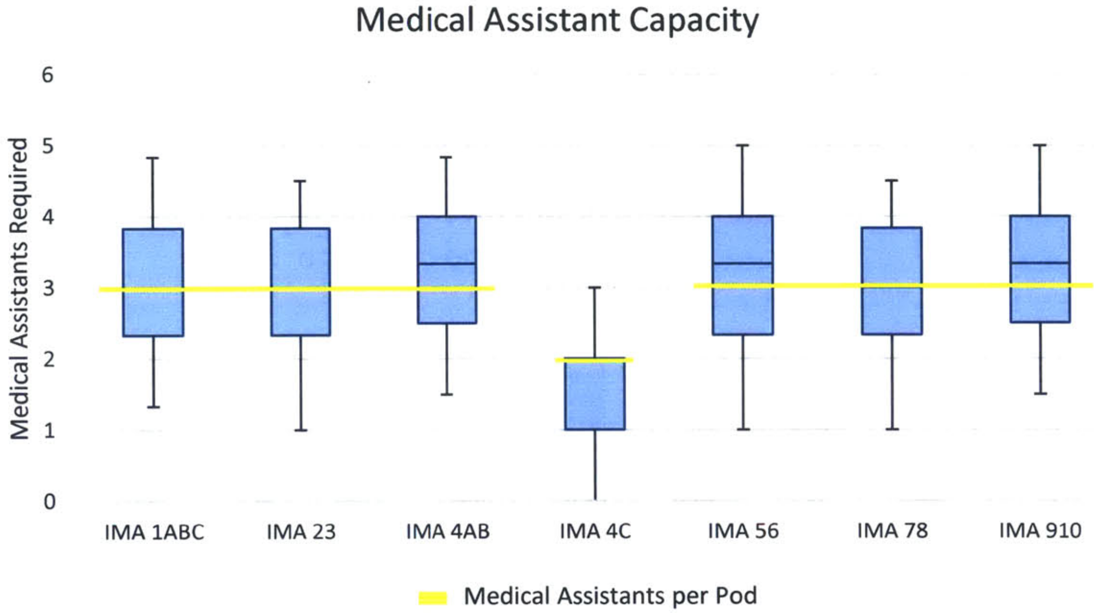


Figure 12 Medical assistants that would be required based on provider presence.

We used daily appointment data to determine the number of active providers in clinical session during all session slots from June 2012 - June 2013. Then, we translate the number of active providers into the number of medical assistants required to support their presence according to the standard IMA staffing ratios. IMA management suggests that one medical assistant can support one attending physician, two nurse practitioners, or three residents. For example, if one attending and two residents are in session, then 1.67 medical assistants are required to support them. Figure 12 shows the distribution of required medical assistants for all pods from June 2012 – June 2013. All pods in the IMA have three medical assistants, except IMA Pod 4C has two.

A majority of clinical sessions in all pods are either over- or under-capacity, while a minimal number are at-capacity. The variability outlined in the sections above translate to non-standard provider schedules week over week where most pods have over-capacity days, except IMA Pod

4C. In aggregate, sufficient medical capacity exists to support the Master Grid Schedule such that an optimal schedule would result in 100% under or at-capacity clinical sessions. Although resident's additional clinical sessions during their ACR block is not accounted for on the Master Grid Schedule, medical assistant capacity can support this workload.

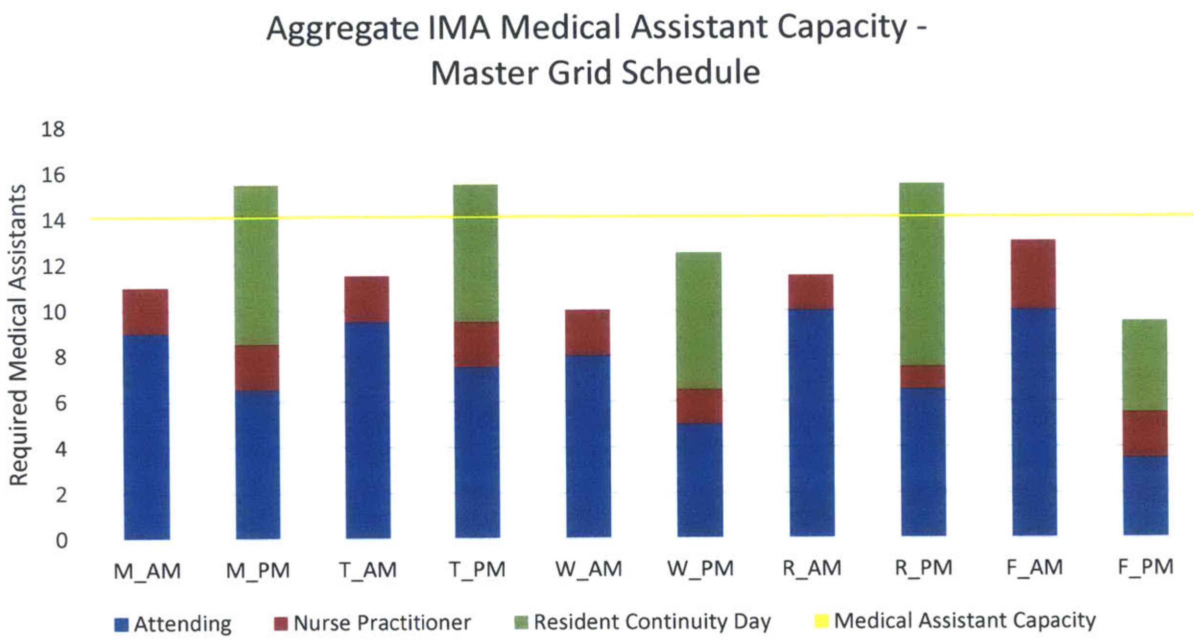


Figure 13 Required number of medical assistants based upon provider scheduling from Master Grid.

Here we translate the aggregate IMA Master Grid Schedule into the number of medical assistants required to support those particular provider sessions per the IMA standard. In sum, there exists adequate medical assistant capacity. However, the unbalanced schedule leave opportunity for improvement to achieve a state where all session slots are at- or under-capacity.

3.5.7. Booked Patient Rates

The data analysis above highlights the sources of variability and quantitative metrics to measure clinical session variability in terms of when they occur. At the same time, there is variability between providers in terms of the number of patients scheduled per hour. Booked patients are those patients that are expected to arrive to the IMA and have scheduled appointments during a clinical session. Individual attending physicians and nurse practitioners have a preferred appointment length for their patients, some typically book 15 minute appointments, while others book for 30 minutes per patient. This individual difference is directly relates to the workload on the particular medical assistants supporting each provider. Residents, on the other hand, do not have this scheduling flexibility. Therefore, we assume that the expected booked patient rate for residents is 2 patients per hour. The following Figure 14 shows the distribution of booked patient rate per hour per provider, whose identity is masked:

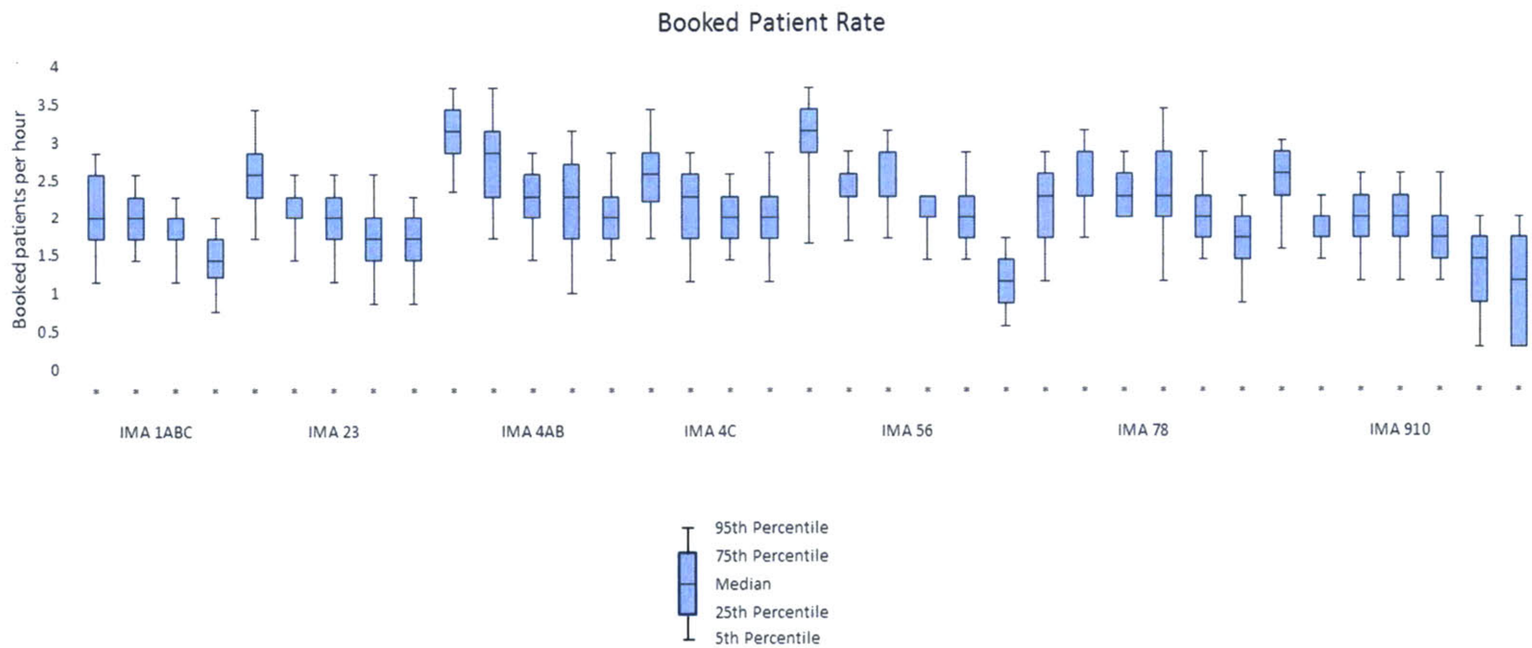


Figure 14 Booked patient rate per attending physician and nurse practitioner.

Appointment level data is used to determine the booked patient rate per provider per clinical session. For every clinical, extra, and make-up session that a provider holds, we use historical appointment data to determine the number of patients that were expected to arrive during that session. Then, given that not all provider sessions are exactly 3.5 hours, we determine the clinic session length by considering the scheduled time of the first and final booked appointments. The booked patient rate can then be approximated by the following Equation 2:

$$Booked\ Patient\ Rate = \frac{Number\ of\ expected\ patient\ appointments}{(Final\ TIme - Start\ Time + 30\ minutes)} \quad (2)$$

We assume that the final appointment is scheduled for 30 minutes. Each provider is identified by a unique code on the x-axis.

Figure 14 shows that the majority of providers schedule 2 – 3 patients per hour in their clinical sessions. However, we note that there is high variability among an individual provider's sessions and across providers. For example, provider #14082 books a median of 2 patients per hour, but data shows this provider does booked as high as 3 and as low as 1 patient per hour.

We characterize the aggregate load on the pod, and therefore the support staff, by the total expected patient load of each provider scheduled per session slot. Combining the results from Figure 14 and the Master Grid Schedule from June 2012 – June 2013, Figure 15 shows the expected patient load per session slot for Pod 2/3 (see Appendix A for other pods).

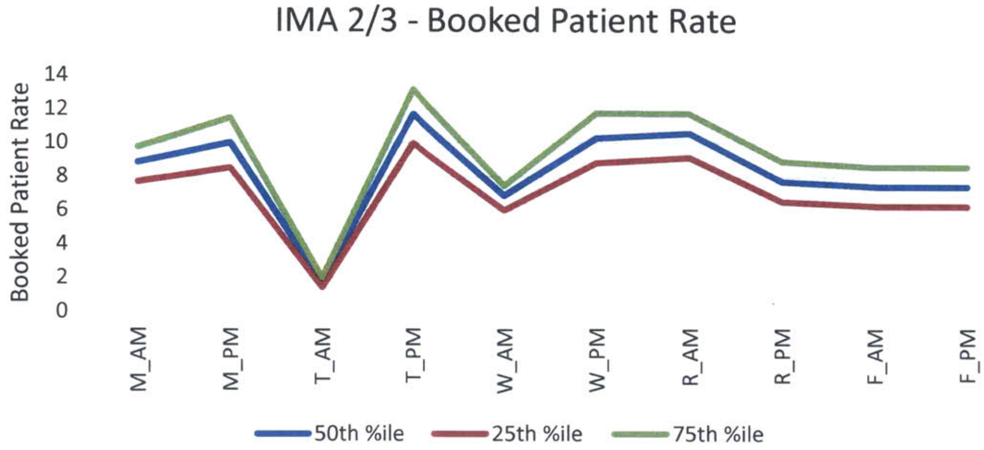


Figure 15 Expected booked patient rate for IMA 2/3 based on the Master Grid Schedule.

The analysis shows the variability in expected patient load per hour over the week. This characterizes the unbalanced workload for support staff.

3.6. Discussion

The data analysis we present in this chapter quantifies the schedule variability along two dimensions, (1) variability in number of concurrent sessions and (2) variability in expected booked patient rates per provider. Figure 15 above highlights how systematic variability present in the Master Grid Schedule along with variability in provider preferences for booked patient rates leads to an overburdened support staff. For Pod 1ABC, for example, the Monday PM slot and Friday AM slot exhibit greater than 5 times the expected patient load than the Wednesday PM slot.

Make-up sessions are scheduled by providers mostly at their discretion, with little consideration to the impacts on support staff. This variability further exacerbates the chaotic nature of the weekly schedule. Additionally, extra sessions create strain on the pod as an unplanned demand on support staff capacity. As noted in Figure 1, providers and MAs work closely to deliver care to patients. Unexpected patient arrivals during a provider's administrative session increases workload leading to burnout among the staff.

This page intentionally left blank.

4. An Improved Provider Scheduling Model

Given the qualitative and quantitative measures of variability discussed, allowing providers to schedule their own clinical sessions in an uncontrolled manner allows for suboptimal week over week schedules to exist. There is an opportunity to create a better system with a significantly more balanced workload on the support staff. The aim of this project is to create a tool that will generate guidelines in scheduling provider clinical sessions that will incorporate a revised definition of workload on support staff into these decisions.

A Revised Medical Assistant Capacity Definition

As described in Section 3.2 and 3.4.3, the IMA management currently considers medical assistant load in terms of the number of concurrent provider sessions that each medical assistant must support. This definition is segmented by provider type, where one medical assistant can support one attending physician, two nurse practitioners, or three residents. However, as noted in Figure 1, the nature of support provided does not vary significantly between the different provider types. A clinical session consumes similar resources from medical assistant regardless of provider type. However, the distinguishing factor depends on the number of expected patient arrivals per hour. The patient booked rate must be incorporated when considering the workload of medical assistants. Therefore, from discussions with the team, we conclude that each medical assistant's workload should be constrained to two concurrent provider sessions and balanced workload measured by the expected booked patient rate. By carefully scheduling clinical sessions, we can balance the expected workload throughout the week in each pod and alleviate stress on the medical care team.

4.1. Optimization Model

We aim to reduce the number of over- and under-capacity days to create a balanced workweek with an optimal provider schedule. We develop an integer linear program to serve as a scheduling tool. It will determine optimal provider schedules that will create a balanced weekly workload for support staff while maintaining system level constraints.

Data Inputs and Outputs

The number of clinical hours that need to be scheduled per provider is determined from the Master Grid Schedule as of December 2014. The booked patient rate is calculated as described in Section 3.5.7. The provider-medical assistant relationship was determined by surveying the pods in December 2014. We recorded the existing working relationships between specific providers and medical assistants. Due to the importance of maintaining these working relationships, providers are scheduled into their existing pod assignments and do not switch to another pod. The preceptor-resident relationship is determined by the residency program from 2014. Each year, as 3rd year residents graduate and incoming 1st year residents arrive, residents are assigned to available preceptors.

The session slots will be similar to the current state, in that each session will be 3.5 hours either in the AM (9AM – 12:30PM) or PM (1:30PM – 5:00PM) throughout the week. Currently, however, many providers have flexible clinic times that fall outside of these standard times, where sessions are shorter than 3.5 hours or are scheduled during the lunch hour. The optimization tool assumes that clinic schedules will conform to the 3.5 hour standard.

The expected output of the model is a binary matrix, 131 rows by 10 columns. Here, we have attending clinics, preceptor slots, nurse practitioner clinics, resident continuity day, and make-up sessions assigned to a particular session slot – all other trivial decision variables are discarded in the output. The following optimization model outlines the mathematical formulation, followed by a qualitative discussion. The optimization model is written in a commercially available software, AMPL using CPLEX 11.2.1 as the solver.

Sets:

$A = set\ of\ attending\ physicians\ (a)$

$A\_2345 = set\ of\ attending\ physicians\ (a)\ that\ hold\ less\ than\ 6\ sessions$

$N = set\ of\ nurse\ practitioners\ (n)$

$R = set\ of\ residents\ (r)$

$M = set\ of\ medical\ assistants\ (m)$

$S = set\ of\ sesions\ (i)$

$SAM = set\ of\ morning\ clinics$

$SPM = set\ of\ afternoon\ clinics$

$S1 = set\ of\ sessions\ on\ Monday, Tuesday$

$S2 = set\ of\ sessions\ on\ Wednesday, Thursday, Friday$

$P = set\ of\ pods\ (p)$

Decision Variables:

$$xs_{api} = \begin{bmatrix} 1 & if\ attending\ (a)\ in\ pod\ (p)\ has\ clinic\ on\ session\ (i) \\ 0 & otherwise \end{bmatrix} \quad (3)$$

$$xp_{api} = \begin{bmatrix} 1 & if\ preceptor\ (a)\ in\ pod\ (p)\ is\ precepting\ on\ session\ (i) \\ 0 & otherwise \end{bmatrix} \quad (4)$$

$$ys_{npi} = \begin{bmatrix} 1 & \textit{if nurse practitioner (n) in pod (p) has clinic on session (i)} \\ 0 & \textit{otherwise} \end{bmatrix} \quad (5)$$

$$zs_{ri} = \begin{bmatrix} 1 & \textit{if resident (r) has continuity clinic on session (i)} \\ 0 & \textit{otherwise} \end{bmatrix} \quad (6)$$

$$w_{pi} = \begin{bmatrix} 1 & \textit{if pod (p) is scheduled for make} - \textit{up session on session (i)} \\ 0 & \textit{otherwise} \end{bmatrix} \quad (7)$$

$$xx_{ap} = \begin{bmatrix} 1 & \textit{if attending (a) is in pod (p)} \\ 0 & \textit{otherwise} \end{bmatrix} \quad (8)$$

$$yy_{np} = \begin{bmatrix} 1 & \textit{if nurse practitioner (n) is in pod (p)} \\ 0 & \textit{otherwise} \end{bmatrix} \quad (9)$$

$$ws1_{ai} = \textit{number of schedule deviations for attending (a)} \quad (10)$$

$$ws2_{ni} = \textit{number of schedule deviations for nurse practitioner (n)} \quad (11)$$

$$wx_{api} = \textit{make} - \textit{up session auxiliary} \quad (12)$$

$$wy_{npi} = \textit{make} - \textit{up session auxiliary} \quad (13)$$

Parameters:

$CSx_{ai} = \textit{current state clinic assignment of attending (a) in session (i)}$
$CSy_{ni} = \textit{current state clinic assignment ofnurse practitioner (p) in session (i)}$
$CPx_{ap} = \textit{current pod assignment ofattending (a) in pod (p)}$
$CPy_{np} = \textit{current pod assignment ofnurse practitioner (p) in pod (p)}$
$BPrate_a = \textit{booked patient rate of attending (a)}$
$BPrate_n = \textit{booked patient rate of nurse practitioner (n)}$
$BPrate_r = \textit{booked patient rate of resident (r)}$
$MAx_{am} = \textit{medical assistant (m) to attending (a) relationship}$
$MAy_{an} = \textit{medical assistant (m) to nurse practitioner (n) relationship}$
$MAz_{ar} = \textit{medical assistant (m) to resident (r) relationship}$
$resPod_{rp} = \textit{resident (r) to pod (p) assignment}$
$resPre_{ra} = \textit{resident (r) to preceptor (a) assignment}$
$attSession_a = \textit{number of clinics per attending (a)}$
$attPrecept_a = \textit{current preceptors sessions (a)}$
$npSession_n = \textit{number of clinics per nurse practitioner (n)}$
$makeups_p = \textit{number of makeup session per pod (p)}$
$rooms_p = \textit{number of exam rooms per pod (p)}$

Objective:

minimize

$$\sum_{p\ in\ P} (MaxLoad_p - MinLoad_p) + \alpha \left(\sum_{a\ in\ A} \sum_{i\ in\ S} ws1_{ai} + \sum_{n\ in\ N} \sum_{i\ in\ S} ws2_{ni} \right) \tag{14}$$

subject to:

$$\sum_{i\ in\ S} xs_{api} = attSession_a * xx_{ap}\ \forall\ a \in A, \forall\ p \in P \tag{15}$$

$$\sum_{i\ in\ S} xp_{api} = attPrecept_a * xx_{ap}\ \forall\ a \in A, \forall\ p \in P \tag{16}$$

$$\sum_{i\ in\ S} ys_{npi} = npSession_n * yy_{np}\ \forall\ n \in N, \forall\ p \in P \tag{17}$$

$$\sum_{i\ in\ S} zs_{ri} = 1\ \forall\ r \in R \tag{18}$$

$$\sum_{i\ in\ S1} xs_{api} \leq floor\left(\frac{attSession_a}{2}\right)\ \forall\ a \in A, \forall\ p \in P \tag{19}$$

$$\sum_{i\ in\ S2} xs_{api} \leq ceil\left(\frac{attSession_a}{2}\right)\ \forall\ a \in A, \forall\ p \in P \tag{20}$$

$$\sum_{i\ in\ SAM} xs_{api} \geq 1 * xx_{ap}\ \forall\ a \in A\_2345, \forall\ p \in P \tag{21}$$

$$\sum_{i\ in\ SPM} xs_{api} \geq 1 * xx_{ap}\ \forall\ a \in A\_2345, \forall\ p \in P \tag{22}$$

$$\sum_{a\ in\ A} 2 * xs_{api} + xp_{api} + \sum_{n\ in\ N} 2 * ys_{npi} + \sum_{r\ in\ R: resPod_{rp}>0} zs_{ri} \leq rooms_p \ \forall\, p \in P, \forall\, i \in S \tag{23}$$

$$\sum_{a\ in\ A} xs_{api} \geq 1 \ \forall\, p \in P, \forall\, i \in S \tag{24}$$

$$\sum_{A\ in\ A: MAx_{am}>0} xs_{api} + \sum_{n\ in\ N: MAy_{nm}} ys_{npi} \leq 2 \ \forall\, p \in P, \forall\, i \in S \tag{25}$$

$$zs_{ri} \geq xp_{api} * resPre_{ra} \ \forall\, a \in A, \forall\, r \in R, \forall\, p \in P, \forall\, i \in S \tag{26}$$

$$Staff_{pi} = \sum_{a\ in\ A} xs_{api} * BPrate_a + \sum_{n\ in\ N} xs_{npi} * BPrate_n + \sum_{r\ in\ R} zs_{ri} * BPrate_r \ \forall\, p \in P, \forall\, i \in S \tag{27}$$

$$Staff_{pi} \leq MaxLoad_p \ \forall\, p \in P, \forall\, i \in S \tag{28}$$

$$Staff_{pi} \geq MinLoad_p \ \forall\, p \in P, \forall\, i \in S \tag{29}$$

$$\sum_{p\ in\ P} xx_{ap} = 1 \ \forall\, a \in A \tag{30}$$

$$\sum_{p\ in\ P} yy_{np} = 1 \ \forall\, n \in N \tag{31}$$

$$xs_{api} + xp_{api} \leq 1 \ \forall\, a \in A, \forall\, p \in P, \forall\, i \in S \tag{32}$$

$$z_{ri} = 0 \ \forall\, r \in R, \forall\, i \in SAM \tag{33}$$

$$ws1_{ai} \geq CSx_{ai} - \sum_{p\ in\ P} xs_{api} \quad \forall\ a\ \in A, \forall\ i \in S \tag{34}$$

$$ws1_{ai} \geq \sum_{p\ in\ P} xs_{api} - CSx_{ai} \quad \forall\ a \in A, \forall\ i\ \in S \tag{35}$$

$$ws2_{ni} \geq CSy_{ni} - \sum_{p\ in\ P} ns_{npi} \quad \forall\ n\ \in N, \forall\ i \in S \tag{36}$$

$$ws2_{ni} \geq \sum_{p\ in\ P} ns_{npi} - CSy_{ni} \quad \forall\ n \in N, \forall\ i\ \in S \tag{37}$$

$$\sum_{i\ in\ S} wx_{api} \geq attSession_a + attPrecept_a + 1\ \forall\ a \in A, \forall\ p \in P \tag{38}$$

$$wx_{api} \geq xs_{api}\ \forall\ a \in A, \forall\ p \in P, \forall\ i \in S \tag{39}$$

$$wx_{api} \geq w_{pi}\ \forall\ a \in A, \forall\ p \in P, \forall\ i \in S \tag{40}$$

$$wx_{api} \geq xp_{pi}\ \forall\ a \in A, \forall\ p \in P, \forall\ i \in S \tag{41}$$

$$\sum_{i\ in\ S} w_{pi} = 2\ \forall\ p \in P, \forall\ i \in S \tag{42}$$

$$wy_{npi} \geq npSession_n + 1\ \forall\ n \in N, \forall\ p \in P, \forall\ i \in S \tag{43}$$

$$wy_{npi} \geq ys_{npi} \forall\ n \in N, \forall\ p \in P, \forall\ i \in S \tag{44}$$

$$wy_{npi} \geq w_{pi} \forall\ n \in N, \forall\ p \in P, \forall\ i \in S \tag{45}$$

$$xs_{ap\text{MonAM}} + xs_{apMonPM} \leq 1\ \forall\ a \in A, \forall\ p \in P \tag{46}$$

$$\vdots$$

$$xs_{apFriAM} + xs_{apFriPM} \leq 1\ \forall\ a \in A, \forall\ p \in P \tag{47}$$

The codification of the model is completed with 5,427 decision variables and 5,176 constraints, and the model determines an optimal solution within 1 minute. The model is run simultaneously for all pods because pods operate independently and we constrain providers to their originally assigned pod.

4.2. Decision Variables

The optimization model sets all the elements of the Master Grid Schedule for each pod, except attending administrative sessions. It determines in which session slot clinical sessions for attending physicians and nurse practitioners, resident continuity days, and attending preceptor slots should occur for each pod throughout the week.

Inevitably, variability will arise week over week requiring make-up sessions to be scheduled. The optimization model will create two dedicated session slots during the week for providers to schedule make-up sessions as the need arises.

4.3. Objective Function

The objective function will balance the expected support staff workload across all session slots throughout the week, while minimizing the number of changes to the current state. The total workload for each session slot is defined by the sum of the median patient arrival rate of each provider scheduled for that session, Equation 27.

The booked patient rate per provider is determined by their historical session data as described in Section 3.5.7. The balancing is achieved by minimizing the sum of the difference between the maximum and minimum expected workload throughout all the session slots for all pods.

In order to strike a balance between the benefits and feasibility of implementation of the proposed schedule, we aim to minimize the number of changes from the current schedule by incorporating a penalty for every deviation. This allows the model to find the best solutions while minimizing disruption to the current schedule. The swap penalty is described by the following equation:

$$Swap\ Penalty = \alpha\left(\sum_{a\ in\ A}\sum_{p\ in\ P} ws1_{ap} + \sum_{n\ in\ N}\sum_{p\ in\ P} ws2_{np}\right) \quad (48)$$

For every provider clinical session that is scheduled outside of their current clinical times, we apply a penalty to the objective. In this way, we identify and schedule the high value changes that will balance the expected workload for support staff, while minimizing schedule changes that will not adequately benefit a balanced workload. The current clinical session times are taken as a snapshot of the Master Grid Schedule in December 2014.

4.3.1. Constraints

There are a number of practical and system level constraints that need to be considered in determining the optimal solution. A detailed discussion is given below:

Maintain Provider Clinics – Equation 15, 17, 18, 30, 31, 32

Currently, each provider commits to hold a specific number of clinical session hours per week and in which session slot. For example, a provider may commit to hold 7 hours of clinical session time per week spread over two session slots (3.5 hours each) to satisfy that provider's patient panel needs. The optimization model aims to maintain the number of clinical sessions and the current pod assignment. All clinical sessions must be scheduled in the same pod and providers must not be double-booked. For example, a provider cannot hold a precepting and clinical session at the same time.

Due to flexibility allowed by the IMA, a provider, for example, may spread their 7 session hours into one 2 hour session and a second 5 hour session. The optimization model assumes all clinical sessions into 3.5 hours. Also, in cases where a provider does not commit to a multiple of 3.5 hour clinic times, the optimization will round up to the nearest integer.

Clinic Distribution – Equation 19, 20, 21, 22, 46, 47

In order to increase patient accessibility, it is important to distribute multiple clinical sessions per provider across the week. Therefore, for providers that hold more than one clinical session, we schedule half their sessions on first half of the week (Monday/Tuesday) and the remaining sessions on the second half (Wednesday/Thursday/Friday). Also, we schedule at least one clinic in the AM and PM. Furthermore, for providers that hold less than 6 clinical sessions, we do not allow for a two clinical sessions to be scheduled on the same day (AM and PM slot).

Room Capacity – Equation 23

Attending physicians utilize two exam rooms during in their clinical session, while nurse practitioners and residents use one. Each pod in the IMA has a specific number of exams rooms available and therefore, the optimization model maintains room capacity for each session slot. All pods have 9 exam rooms, except Pod 4C, which has 6 exam rooms.

Medical Assistant Capacity – Equation 25

As noted in Section 3.4.3, individual medical assistants and providers have developed a strong working relationship over the years. In order to preserve the team structure and limit the number of concurrent sessions a medical assistant must support, no more than two providers that work with a particular medical assistant can be scheduled into the same session slot. Thus, a medical assistant supports a maximum of two concurrent clinical sessions.

The provider-medical assistant relationship was taken as a snapshot in December 2014. With the dynamic nature of the practice, there is inherent attrition and new hires, for both providers and staff, into the practice that requires on-going adaptability. However, for the purposes of the model, we take the staffing relationships in December 2014.

Preceptor-Resident Relationship – Equation 16, 26

Individual residents work closely with their assigned preceptor during their three year residency program. Therefore, the preceptor-resident relationship must be maintained by scheduling preceptor slots and resident continuity sessions concurrently.

Resident Continuity Session – Equation 18, 33

Residents hold continuity sessions every week. These sessions are restricted to the afternoon session slots due to their other educational commitments.

Make-Up Session – Equation 38, 39, 40, 41, 42, 43, 44, 45

We allocate two make-up session slots during the week for each pod for when providers need to reschedule clinical sessions. The two make-up session slots are chosen such that at least one slot is outside each provider's standard schedule. In this way, when the need arises to schedule a make-up session, there exists at least one slot during the week to schedule a new session.

Track Deviations – Equations 34, 35, 36, 37

In order to limit the number of changes made to the schedule, the model tracks the number of changes of the proposed solution to the current state. These constraints track the changes via auxiliary variables that are used in the Swap Penalty, described below.

4.4. Results

We present detailed results of the optimization model through a case study on IMA Pod 2/3, remaining pods are shown in Appendix A. For the purposes of this project, IMA Pod 2/3 is a representative pod within the IMA for detailed analysis. First, we examine the current state and then, compare to the optimization model results. Finally, we present aggregate statistics on all IMA pods.

4.4.1. Case study on Pod 2/3

There are 6 attending physicians, 1 nurse practitioner, and 11 residents assigned to Pod 2/3 that are supported by 3 full-time medical assistants. In aggregate, these providers are scheduled to hold 41 clinical sessions throughout the week. In Figure 16 below, we show the Master Grid Schedule for Pod 2/3 in the current state.

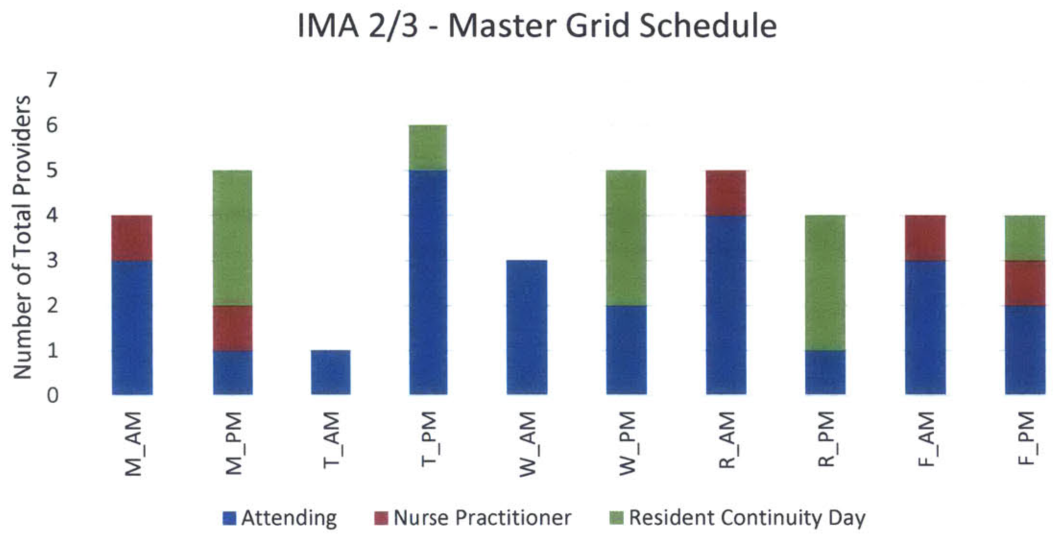


Figure 16 Current state Master Grid Schedule for Pod 2/3.

There is significant variability session over session, for example, the Tuesday AM slot has one scheduled attending while Tuesday PM has five scheduled attending clinical sessions. As evident, systematic variability is built into the schedule causes such significant fundamental differences in session over session workload. Also, either one or three resident continuity days are scheduled per PM session slot. Given the unique workflows and support required by resident physicians, this adds a further layer of complexity for medical assistants.

The IMA also allows attending physicians and nurse practitioners to schedule non-standard session lengths and start or end times. In the following Figure 17, we note the variability of providers (excluding residents) in Pod 2/3 in this respect:

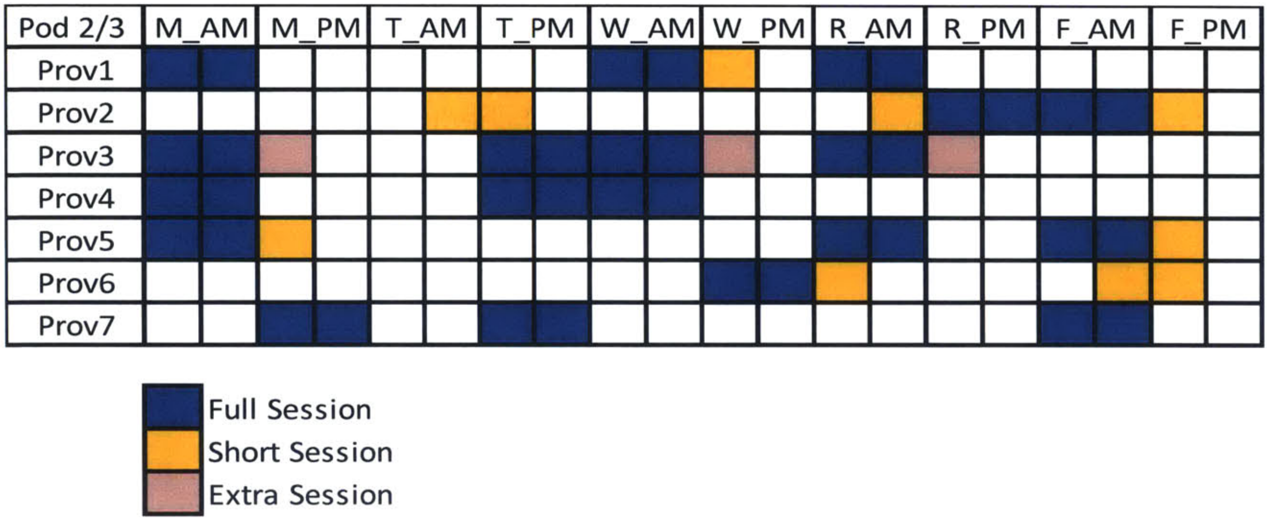


Figure 17 Pod 2/3 schedule by provider outlining full, short, and extra sessions.

Full sessions denote clinical sessions that adhere to the IMA standard of 3.5 hours. Conversely, short sessions are those that are less than 3.5 hours and extra sessions, described in detail in Section 3.5.3, are those short sessions that regularly occur, but are not scheduled on the Master Grid Schedule. In the instances where a short or extra session neighbors a full session, the provider is scheduling patients during the support staff's lunch hour. For example, Prov3 regularly schedules patients into the Monday PM slot, but is only scheduled for a full session Monday AM. Figure 17 highlights the extensive variability in provider session lengths that impose stress onto the support staff. Of the 7 non-resident providers, two providers adhere to the IMA standard. Residents are always scheduled for standard clinical sessions.

It is important to consider the expected booked patient rate associated with the Master Grid Schedule for the pod. Figure 18 outlines the sum of individual provider's booked patient rate for their respective session slots.

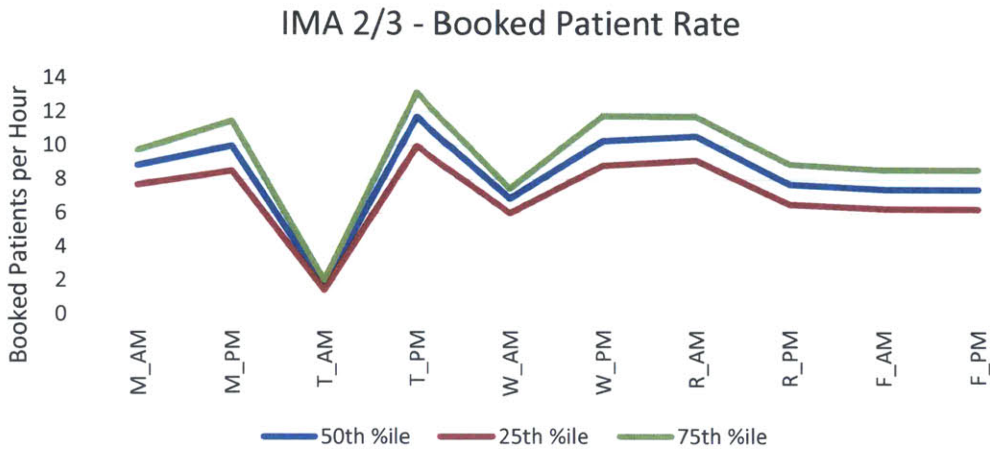


Figure 18 The total expected booked patient rate for Pod 2/3 based upon the current state.

Individual provider patient booking rates, discussed in Section 3.5.7, are tabulated with their clinical session times from the Master Grid Schedule to determine the total expected patients for the pod per session slot. There is a corresponding variability effect on expected patient arrivals, where Tuesday AM has the least workload and Tuesday PM has the most. This creates stress on the support staff and adds to the sense of burnout. Sources of unpredictable variability can further exacerbate the issue when providers schedule make-up sessions without considering medical assistant workload.

The α Parameter

The current state, described above, is used as an input to the optimization model to determine a more balanced schedule. The 41 clinical sessions are optimally rearranged while minimizing the number of deviations to create a more balanced workload for support staff. The swap penalty is used to minimize these deviations, but requires fixing the α parameter to an appropriate value that balances the benefits and deviations of the proposed schedule. For example, the model could achieve the most balanced workload by freely shifting schedules of all providers. However, this is an impractical and unnecessary solution. First, this would allow for redundant shifts and successfully implementing such a proposed schedule is highly unlikely with practical constraints. Therefore, it is important to consider various proposed schedules and the associated magnitude of change by varying the α parameter that is used across all pods. Figure 19 below outlines the results of various scenarios.

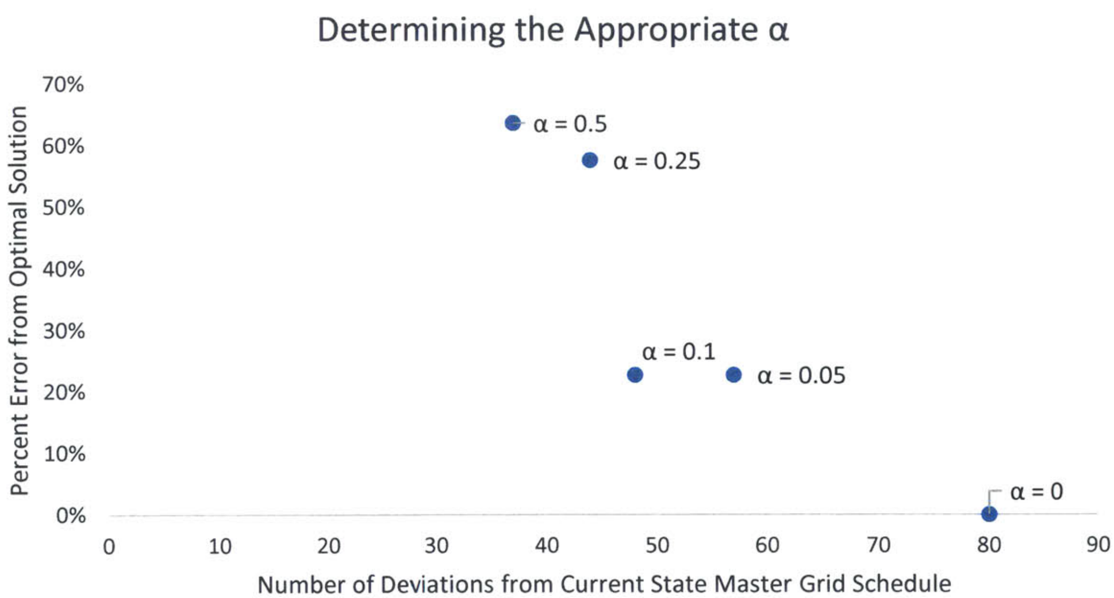


Figure 19 Tradeoffs between optimization solution and number of schedule deviations.

The model determines the optimal solution when $\alpha = 0$, where there are no limitations on the number of deviations, and finds the objective function to be 18.85 requiring 80 provider clinic schedule changes. This is the best-case scenario. For all $\alpha > 0$ cases, we graph the percent difference between the objective function to the optimal objective value for varying values of α against the number of proposed schedule deviations from the Master Grid Schedule. For example, when swaps are penalized heavily, $\alpha = 0.5$, the objective value deviates from the optimal solution by 64%, but requires only 37 provider clinic changes. The percent error is calculated by the difference between maximum and minimum booked patient rates and the additional swap penalty is removed. We determine that to balance schedule changes and maintain a reasonable objective value, we use $\alpha = 0.1$ for the analysis presented. Figure 20 and Figure 21 describe the results of the model:

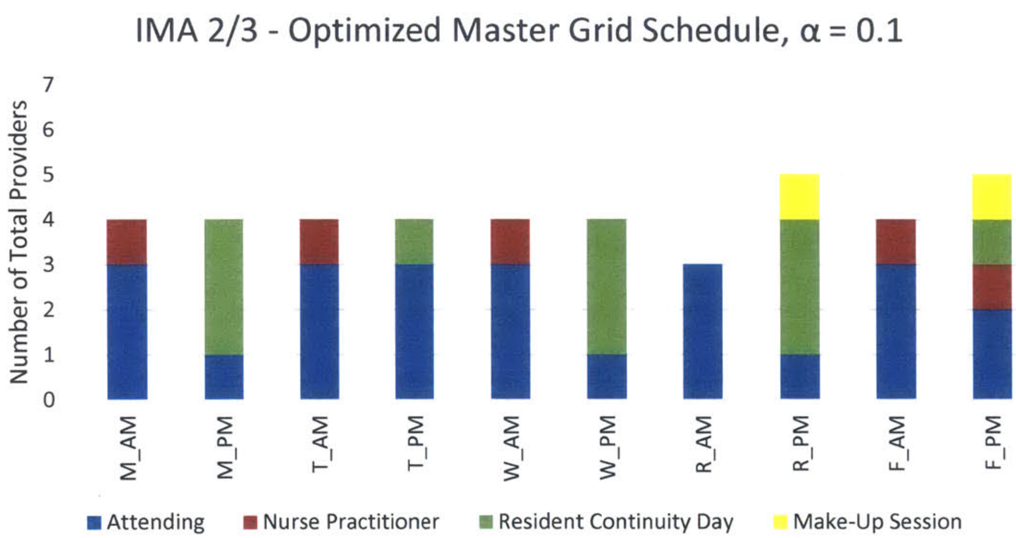


Figure 20 Optimized Master Grid Schedule for Pod 2/3.

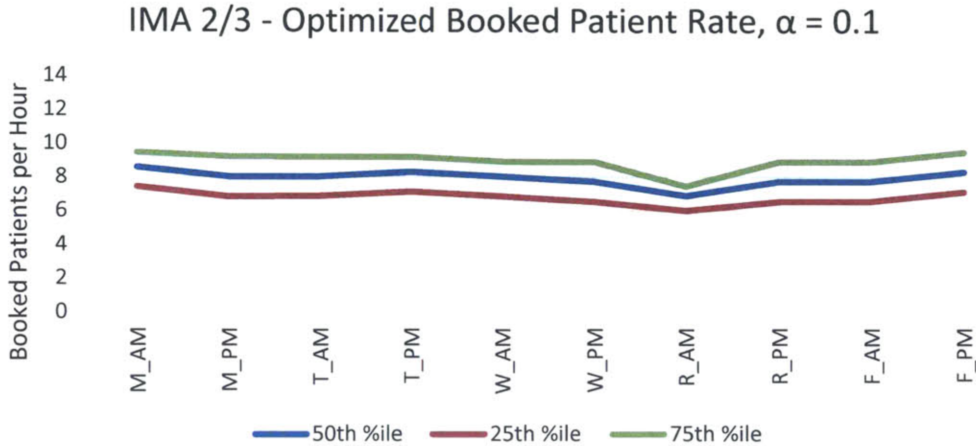


Figure 21 Booked patient rate for Pod 2/3 based on Master Grid Schedule.

Significant improvements are made in balancing the expected number of patients per hour across the day (from AM to PM) and day over day. The degree of variation in expected patients per hour is approximated by the maximum load less minimum load over the week. In the current state, the workload is 10 patients per hour, while the optimized schedule results in a difference of 1.7 – an 83% improvement. The model suggests that 6 attending clinical sessions and 2 nurse practitioner sessions be rescheduled for another session slot to achieve this result – 19.5% of total clinical sessions in Pod 2/3. In the following Figure 22, we compare the $\alpha = 0.1$ case to the $\alpha = 0$ case to examine the effects of limiting deviations:

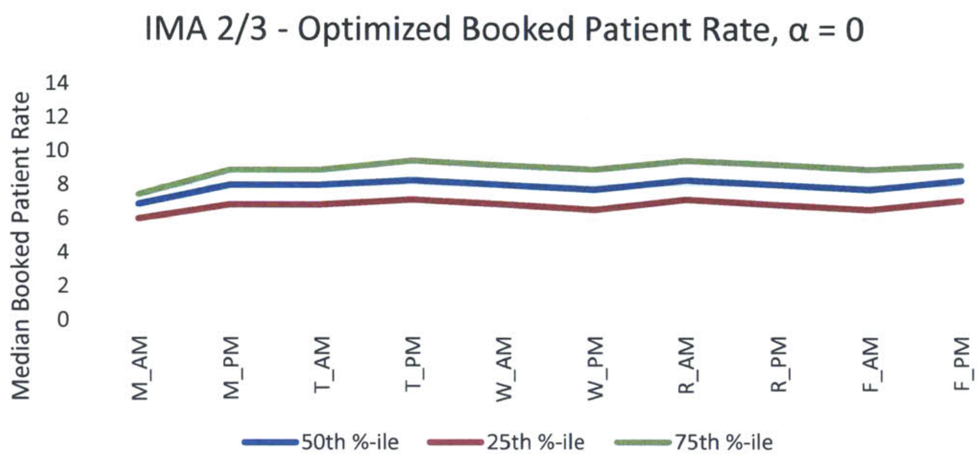


Figure 22 Booked patient rate for Pod 2/3 based on optimized solution with unlimited swaps.

The objective value in this case is 1.429 requiring changes to 11 attending clinics and 2 nurse practitioner sessions. The marginal improvement in booked patient rate is 1.04 per clinic requiring rescheduling (8.3/8) in the $\alpha = 0.1$ case. However, the additional improvement in the $\alpha = 0$ case in patient rate is 0.271 at the cost of 5 additional rescheduled clinics. We determine the added benefit in variability reduction is not worth the added cost of more deviations.

The following Table 1 gives the detailed schedule changes proposed by the model.

	M_AM	M_PM	T_AM	T_PM	W_AM	W_PM	R_AM	R_PM	F_AM	F_PM
Attending 1	0	1	0	1	0	0	*1*	0	1	0
Attending 1	0	1	0	1	0	0	1	0	1	0
Attending 2	0	0	0	1	0	1	0	0	*1*	*1*
Attending 2	0	0	1	0	0	1	0	0	0	1
Attending 3	1	0	0	1	1	0	0	0	0	0
Attending 3	0	0	0	1	1	0	0	0	1	0
Attending 4	1	0	0	1	1	0	1	0	0	0
Attending 4	1	0	0	1	1	0	1	0	0	0
Attending 5	0	0	*1*	*1*	0	0	*1*	1	1	*1*
Attending 5	1	0	1	0	1	0	0	1	1	0
Attending 6	1	0	0	0	1	*1*	1	0	0	0
Attending 6	1	0	1	0	0	0	1	0	0	1
NP 1	1	*1*	0	0	0	0	1	0	1	*1*
NP 1	1	0	1	0	1	0	0	0	1	1

Current State
Optimal Solution

Table 1 Comparison of current state and optimized clinic schedules for individual providers in Pod 2/3.

The current state includes standard and short sessions, denoted by italics, by providers. Additionally, dedicated make-up sessions are scheduled in the optimized state to allocate capacity for week over week variation. When clinical sessions need to be rescheduled, preference is given to these particular session slots.

4.4.2. IMA Statistics

Table 2 below shows the results of the optimization model for the whole practice segmented by pod and by provider type. The table quantifies the number of required clinical session schedule changes to achieve a more balanced workweek. For example, in Pod 1ABC, there were 2 attendings requiring 0 changes, 3 attendings requiring 1 change, and 2 attendings requiring 2 changes each.

	Attending			**Nurse Practitioner**		
Number of Changes	0	1	2	0	1	2
Pod						
IMA 1ABC	2	3	2	0	1	0
IMA 23	2	2	2	0	0	1
IMA 4AB	3	3	1	0	2	0
IMA 4C	1	2	1	0	0	0
IMA 56	0	5	2	0	1	0
IMA 78	2	3	2	0	1	0
IMA 910	1	4	2	0	0	1
Sum	**11**	**22**	**24**	**0**	**5**	**4**

Table 2 Magnitude of change required to achieve proposed schedule.

Table 3 below outlines variability improvement per pod from the current to optimized state and the magnitude of change suggested by the optimization model. Variability is defined as the difference between maximum and minimum expected booked patient rate during the workweek. Magnitude of change is defined as the percentage of clinical sessions that deviate from the current state.

Pod	Variability Current State (patients per hour)	Variability Optimized (patients per hour)	Improvement	Magnitude of Change
IMA 1ABC	8.00	4.00	50.0%	26.7%
IMA 2/3	10.00	1.71	82.9%	19.5%
IMA 4AB	6.29	3.14	50.0%	14.6%
IMA 4C	4.57	2.57	43.8%	28.6%
IMA 5/6	13.14	4.00	69.6%	26.2%
IMA 7/8	8.29	3.43	58.6%	21.1%
IMA 9/10	9.71	4.00	58.8%	26.3%

Table 3 Comparison of variability in aggregate expected booked patient rate of current state and proposed solution.

For all pods, optimally rescheduling 15-25% of clinical sessions results in a 50-60% improvement in workload variability. A small number of high value changes have an outsized effect on creating a balanced schedule. It is noteworthy that Pod 4C showed the lowest improvement for the highest magnitude of change. This is due to the size of Pod 4C; it is the smallest pod with fewest providers. In addition to variability improvement, we consider the impact on the number of concurrent sessions translated into the required medical assistants to support their presence. Figure 16 outlined the current state Master Grid Schedule, using the standard IMA provider-medical assistant support ratios. However, we have shown that provider type is not a distinguishing factor in considering medical assistant capacity, but rather, medical assistants can support a maximum of two concurrent sessions. In the following Figure 23, we show the current aggregate IMA medical assistant capacity using the new definition and in Figure 24, we show the optimized solution.

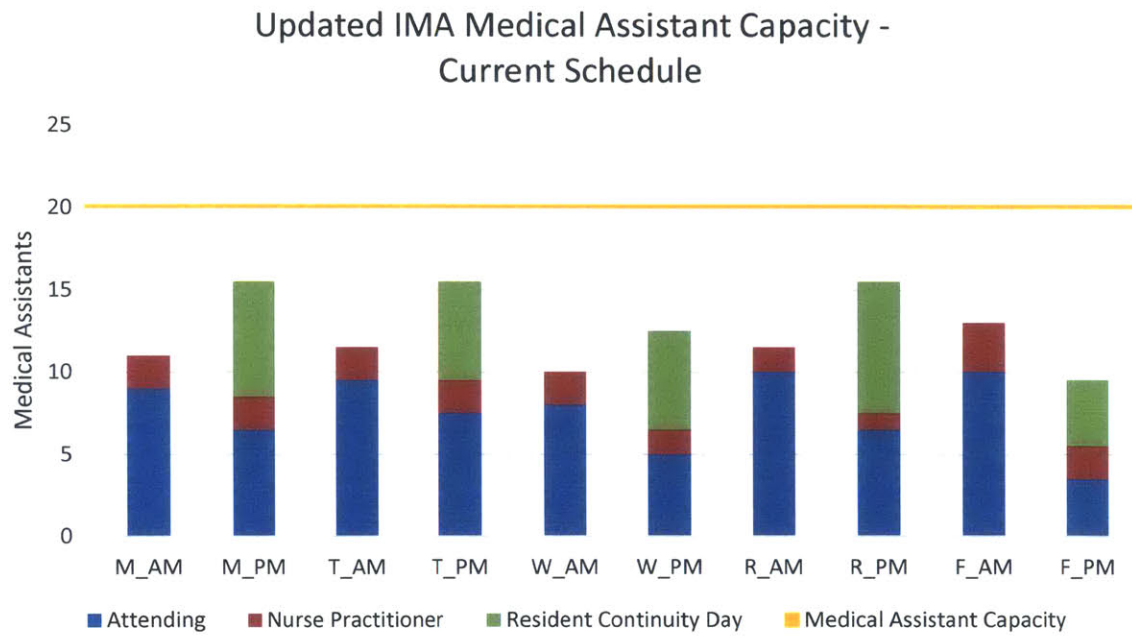


Figure 23 Number of medical assistants required for the given provider presence based on updated support ratios.

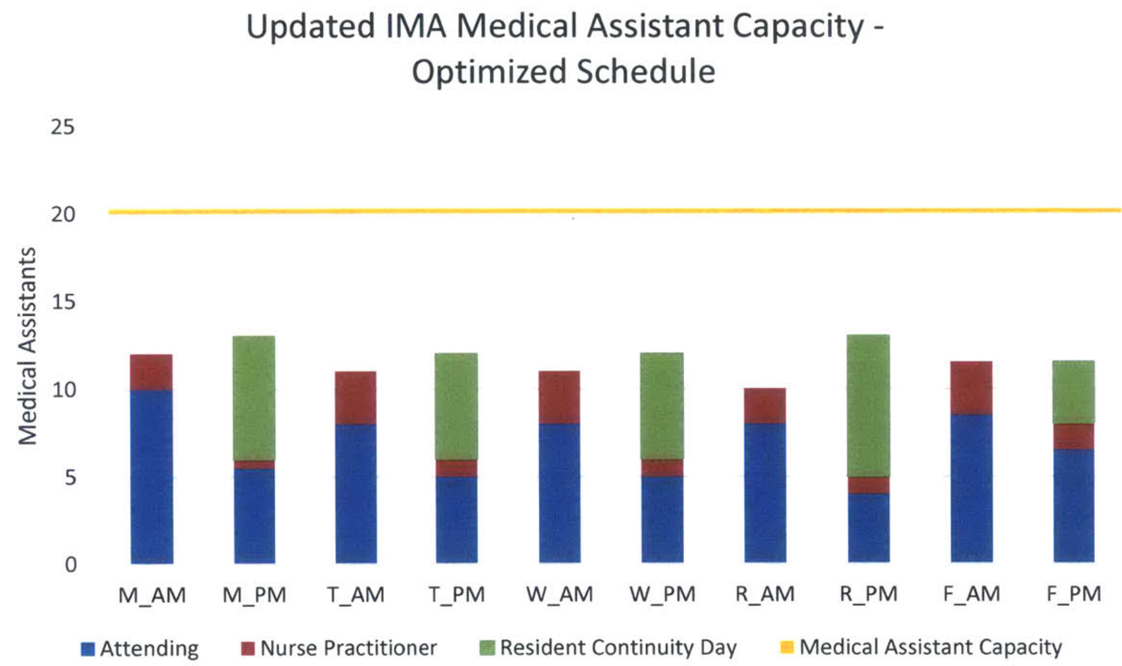


Figure 24 Number of medical assistants required for the given provider presence for the proposed solution.

With an updated definition, we see significant variability in provider presence in the current state. The requirement to have resident continuity days in the PM slot creates unbalance as well as an aversion to scheduling on Friday PM. However, the optimized solution makes high value changes to result in a more balanced provider presence.

5. Conclusion and Recommendations

With strategic changes, the IMA can achieve significant improvements in reducing variation which can lead to enhanced provider-staff relationships and improved patient experience. By addressing the sources of variability, the IMA can proactively manage workload to reduce times where support staff experience overburdened demand. The optimization model can serve as a tool in aiding future scheduling decisions iteratively as new providers or support staff are added to the organization. For example, at the beginning of every academic year, there is a resident changeover where 3rd year residents graduate and leave the program and incoming residents must be scheduled. The model can also suggest when clinics should be scheduled for new providers that join the IMA.

Systematic Variability

There are two primary aspects of systematic variability that the presented work aims to address: (1) provider-determined clinical schedules and (2) variable clinical session length. When each provider schedules based on their particular preferences, it can lead to an unbalanced load from an aggregate perspective of each pod. We recommend that IMA leadership take an active role in determining the Master Grid Schedule per pod to ensure a balanced workload for support staff, rather than allowing providers to self-select. The optimization model serves as a tool in the decision-making process.

The IMA also allows flexibility in clinical session length and start time, which further stresses the support staff. We recommend that each clinical session adhere to the 3.5 hour standard and begin at the prescribed intervals – AM sessions should be 9:00AM–12:30PM and PM sessions should be 1:30PM–5:00PM. This will enhance the support staff experience in that it maintains consistency throughout the session by removing variable session start times. Also, in adopting stricter rules, the support staff lunch hour is protected further improving the support staff experience.

Predictable + Unpredictable

Predictable and unpredictable sources of variability that ultimately arise cause deviations from an ideal Master Grid Schedule on a week over week basis. Currently, make-up sessions are also provider-determined assuming exam room capacity is available. We recommend that IMA management take a proactive role in addressing this need. The optimization model allocates two make-up session slots that providers can be scheduled into as needed. In this way, unexpected make-up session consistently are scheduled on these two session times, which makes the workload more predictable for support staff.

5.1. Implementation Challenges

There are a number of challenges that may arise in implementing the suggested changes. First, there will be organizational inertia in adopting new schedules. A number of providers have been working in the IMA for many years with a regular schedule year over year that may pose difficult to change. These individual preferences must be taken into account. Additionally, providers have other obligations within MGH, such as research and teaching, which may conflict with schedule changes. However, for new providers entering the IMA, management should proactively use the optimization model to schedule their clinics.

Operationally managing make-up session can also pose challenges. Currently, make-up sessions are scheduled based on the individual provider's preferences as long as there is available exam room capacity during that time. We have discussed the unpredictable stresses this puts onto the support staff workload and therefore, we suggest that make-up session be scheduled during predictable session slots. The model places two make-up sessions throughout the week for each pod such that at least one session is available outside a provider's standard clinic schedule. IMA management should proactively leverage these times to schedule make-up sessions.

Appendix A

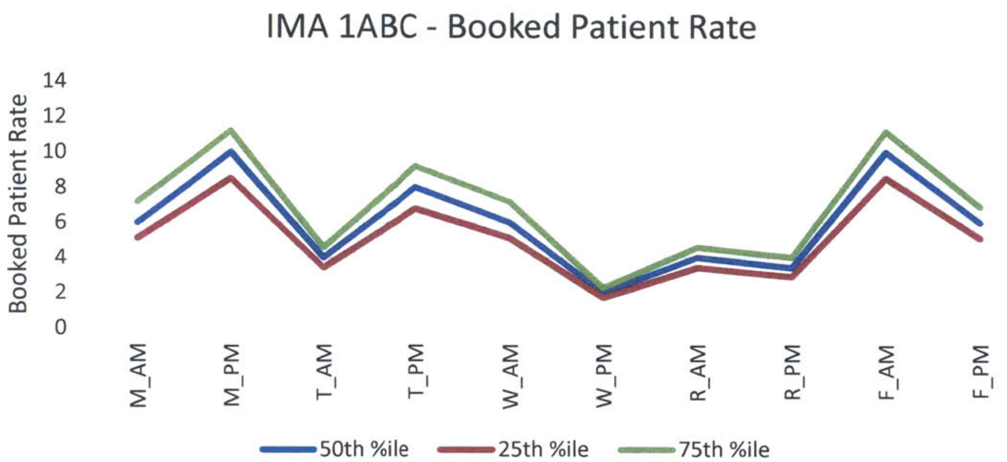


Figure 25 Expected booked patient rate for IMA 1ABC based on current state Master Grid Schedule.

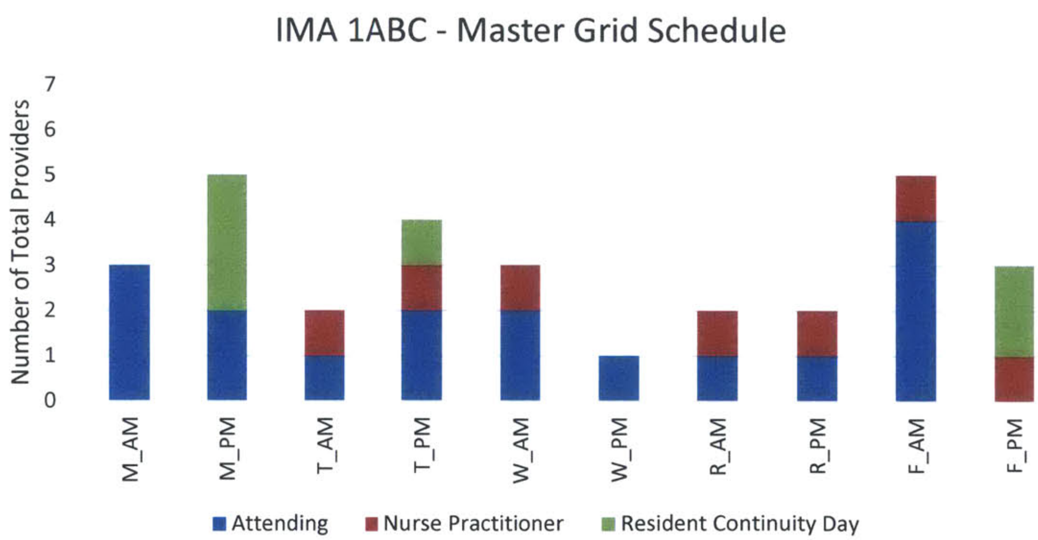


Figure 26 Current state Master Grid Schedule for IMA 1ABC.

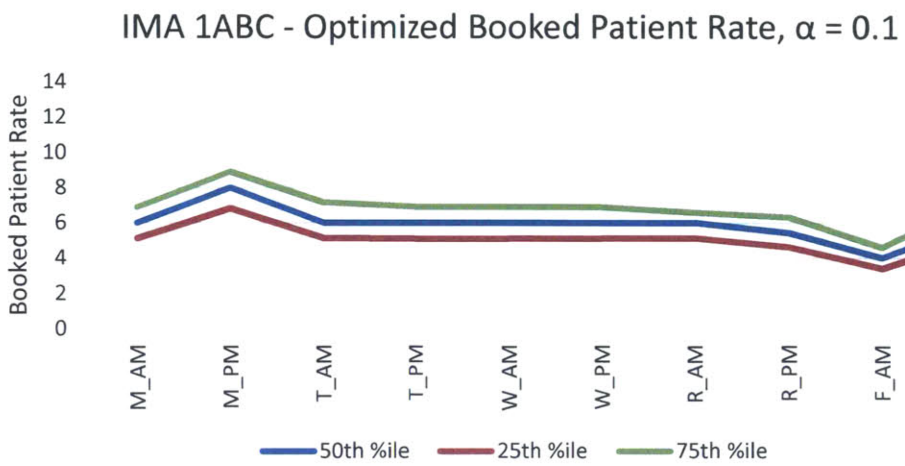


Figure 27 Expected booked patient rate for IMA 1ABC based on optimized Master Grid Schedule.

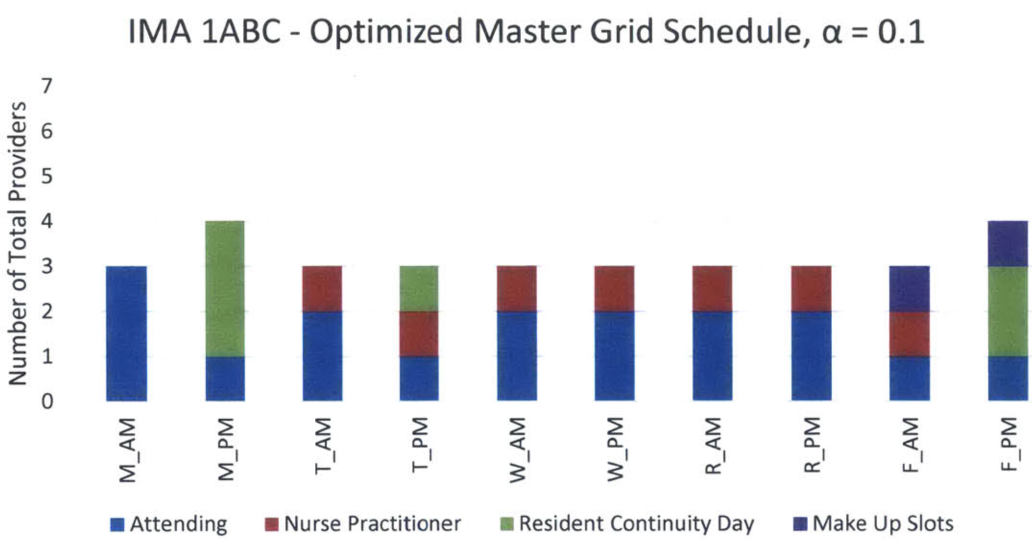


Figure 28 Optimized Master Grid Schedule for IMA 1ABC.

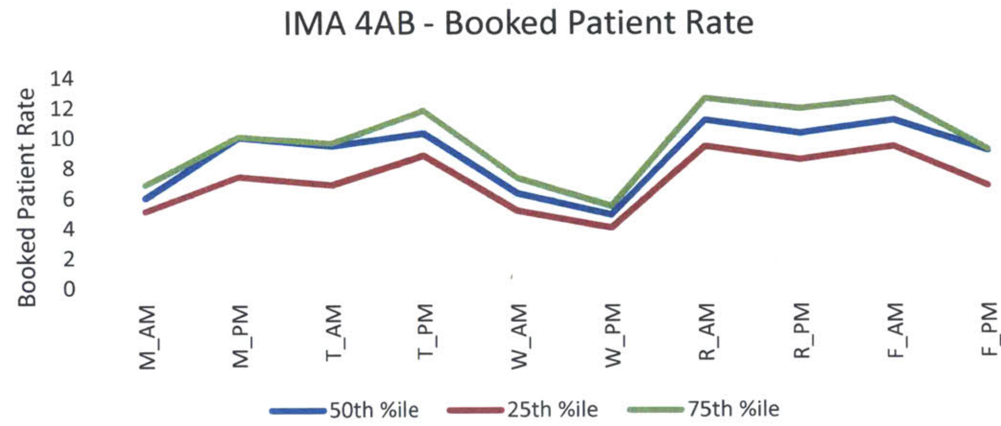


Figure 29 Expected booked patient rate for IMA 4AB based on current state Master Grid Schedule.

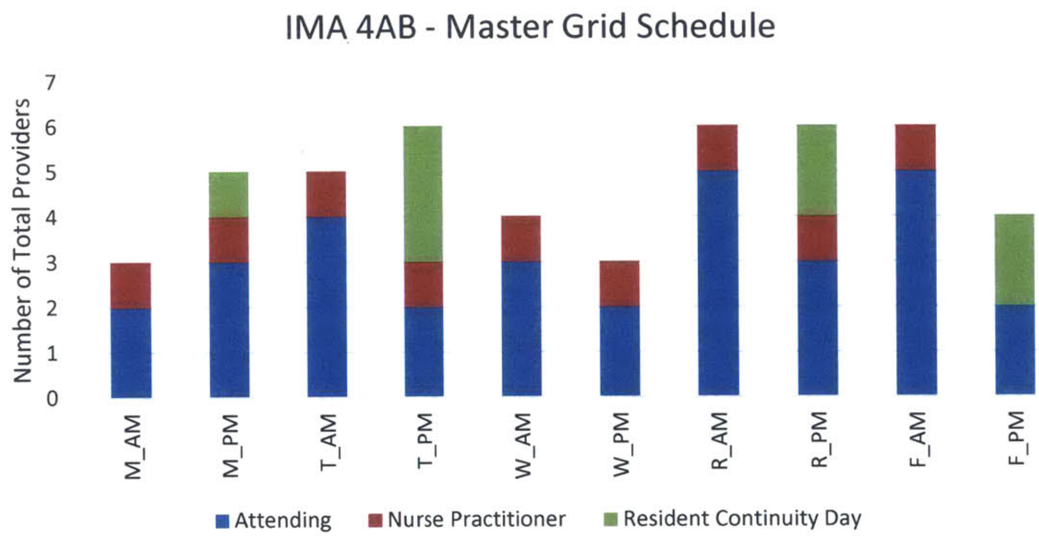


Figure 30 Current state Master Grid Schedule for IMA 4AB

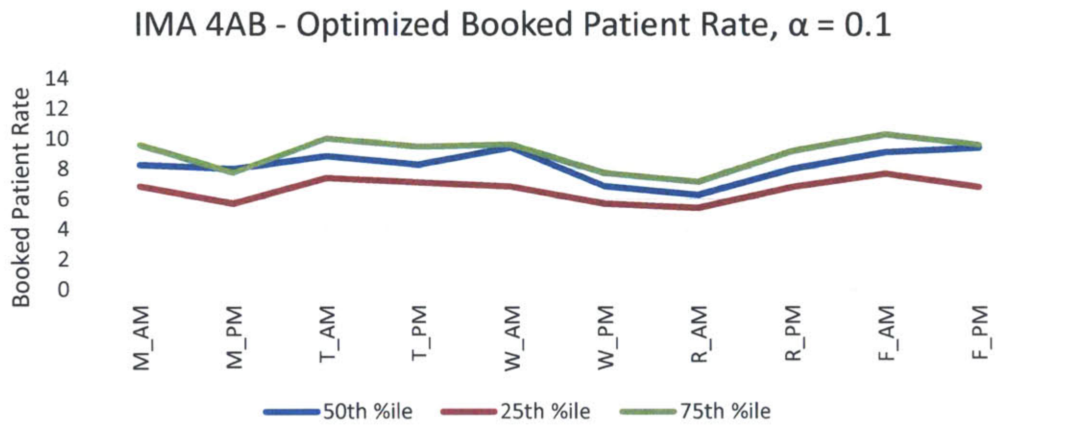


Figure 31 Expected booked patient rate for IMA 4AB based on optimized Master Grid Schedule.

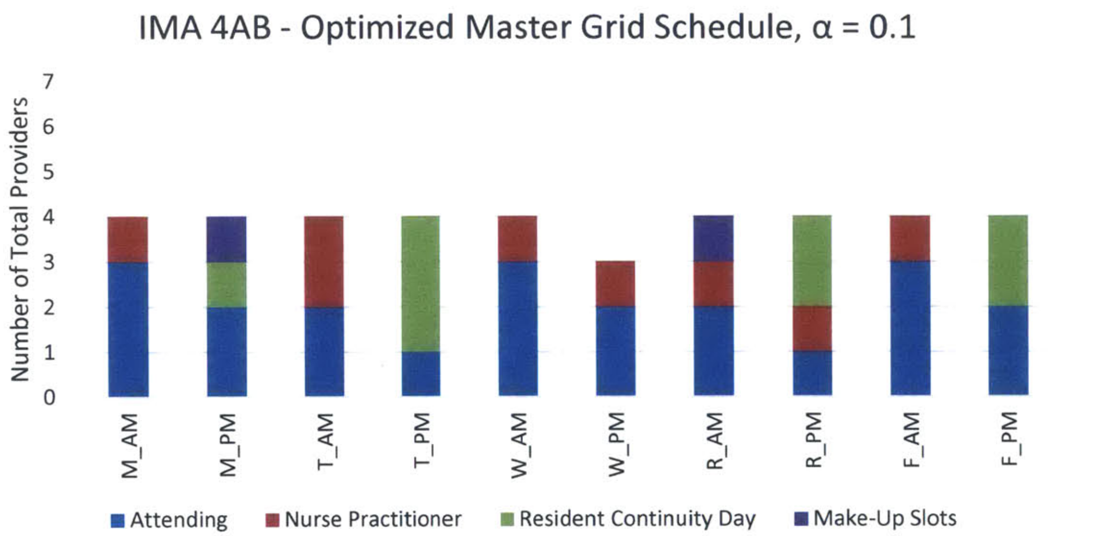


Figure 32 Optimized Master Grid Schedule for IMA 4AB.

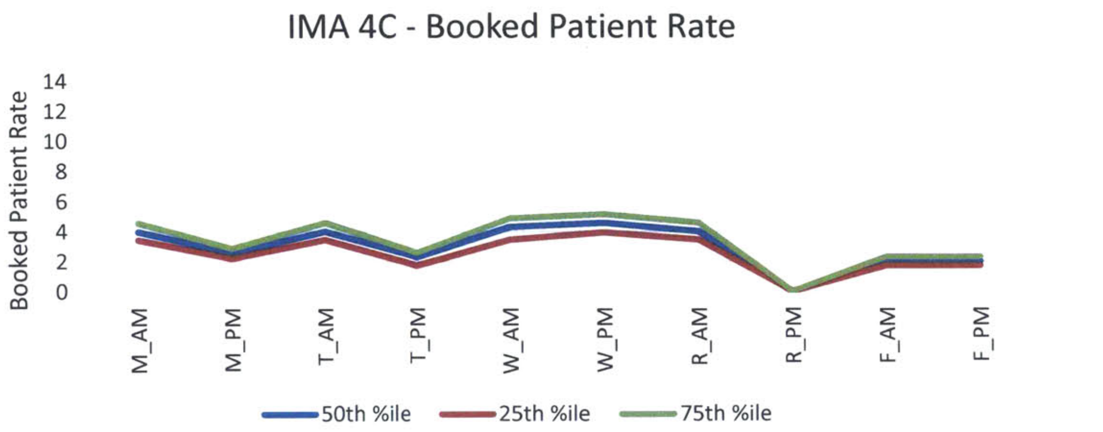


Figure 33 Expected booked patient rate for IMA 4C based on current state Master Grid Schedule.

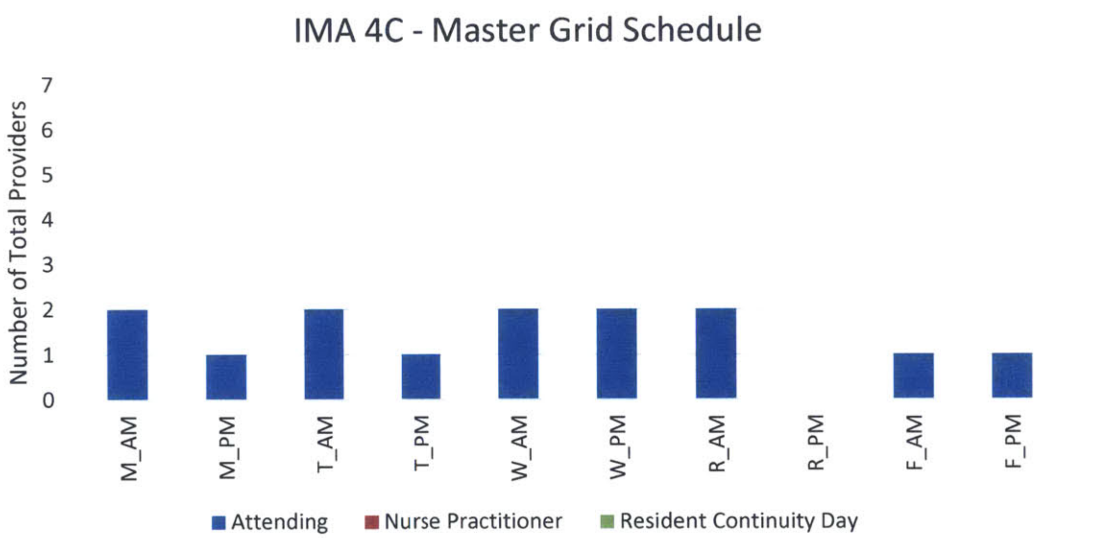


Figure 34 Current state Master Grid Schedule for IMA 4C.

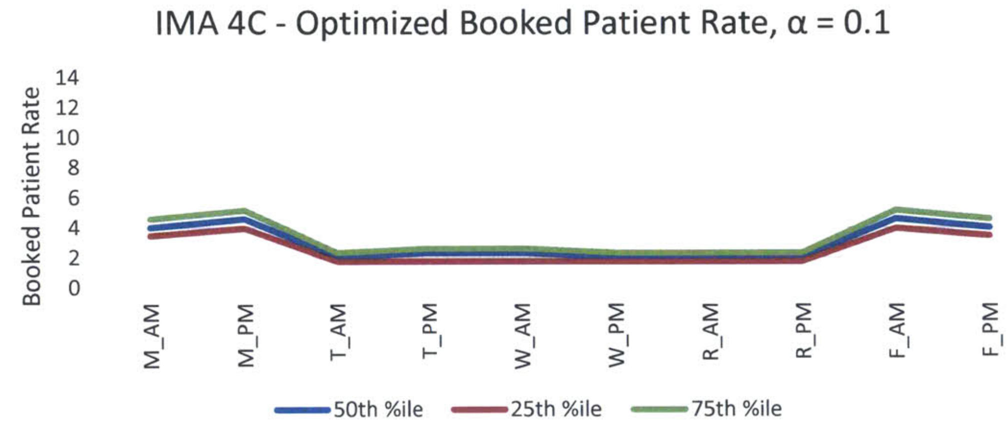


Figure 35 Expected booked patient rate for IMA 4C based on optimized Master Grid Schedule.

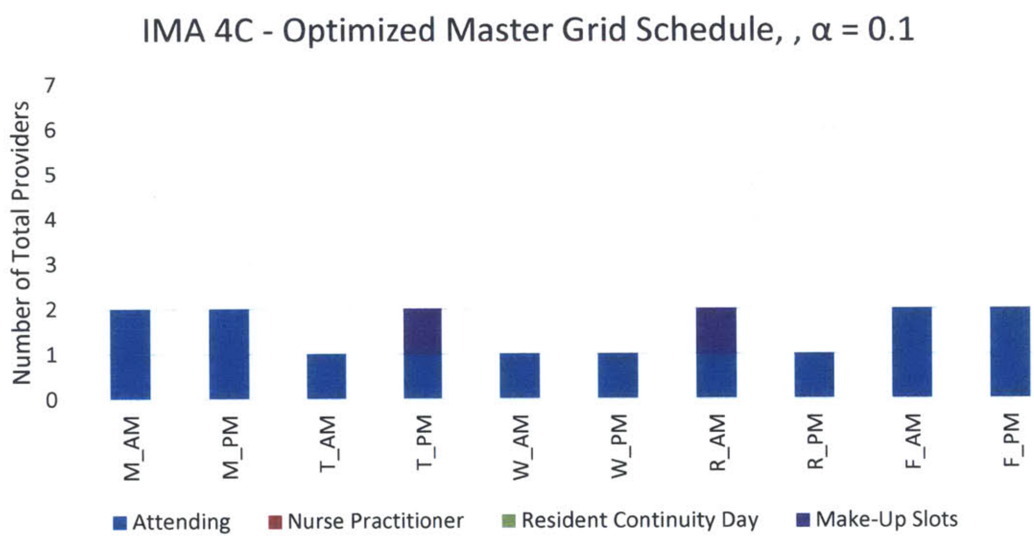


Figure 36 Optimized Master Grid Schedule for IMA 4C.

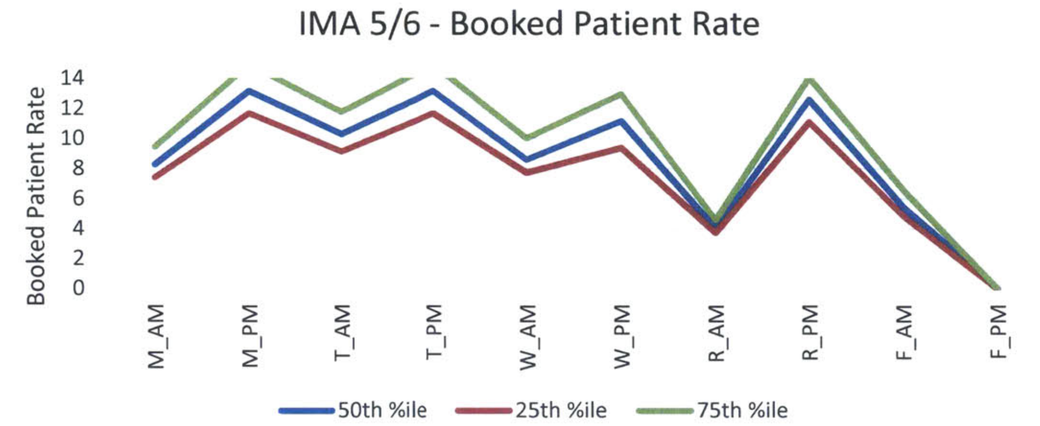


Figure 37 Expected booked patient rate for IMA 5/6 based on current state Master Grid Schedule.

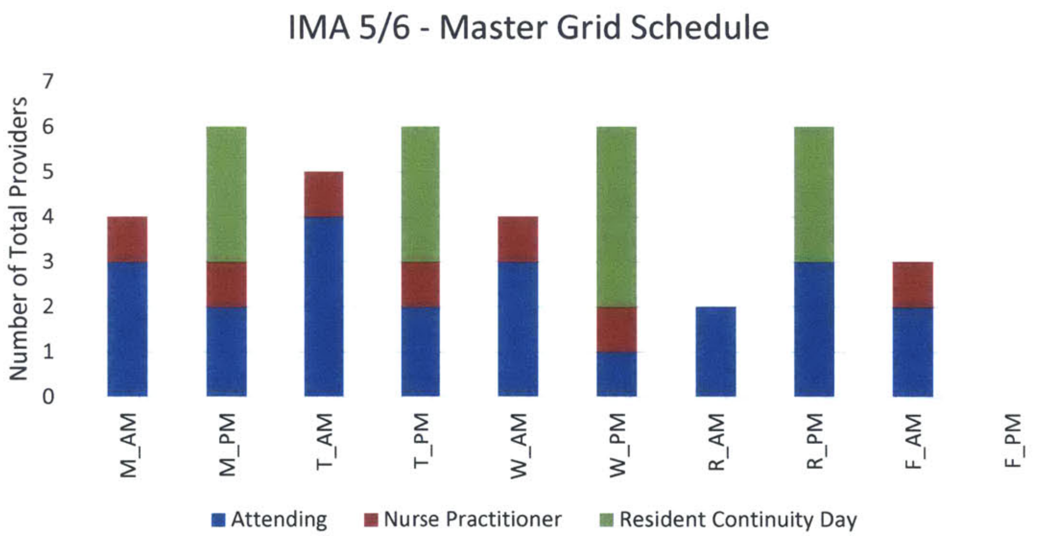


Figure 38 Current state Master Grid Schedule for IMA 5/6.

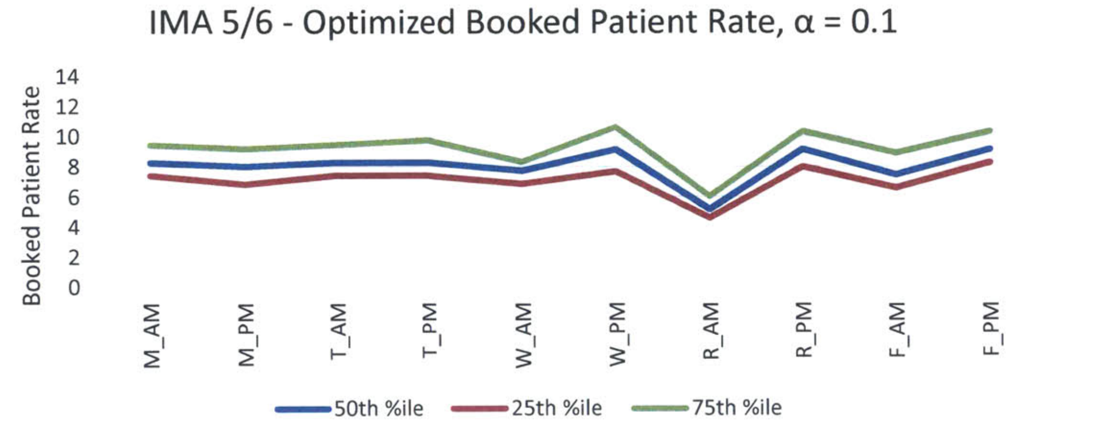


Figure 39 Expected booked patient rate for IMA 5/6 based on optimized Master Grid Schedule.

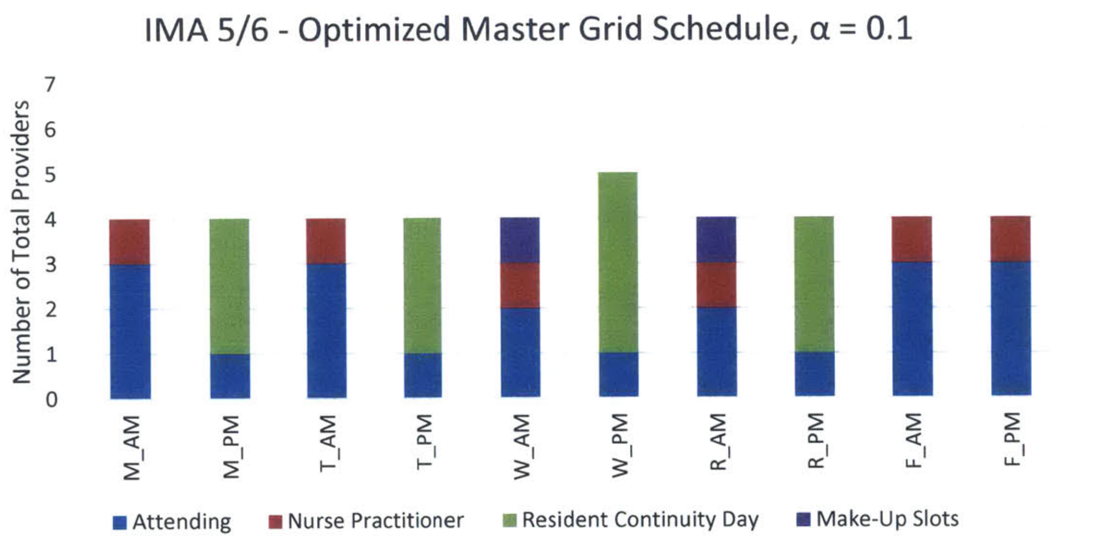


Figure 40 Optimized Master Grid Schedule for IMA 5/6.

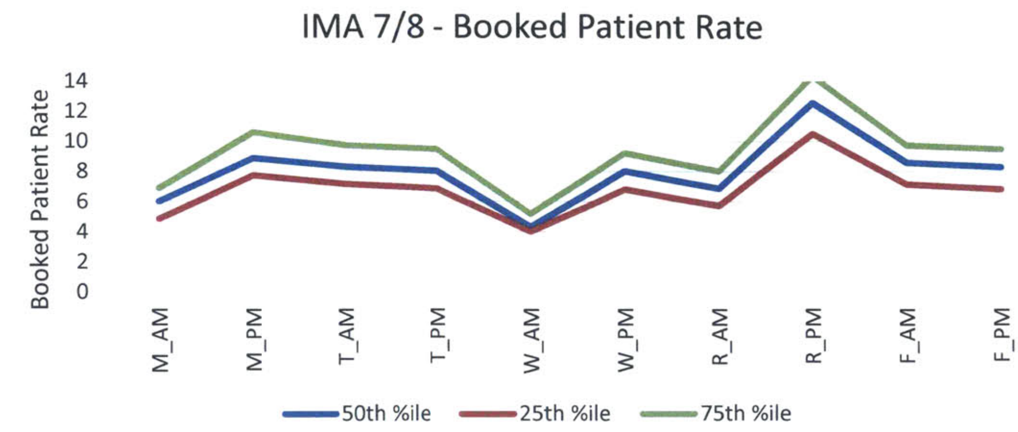


Figure 41 Expected booked patient rate for IMA 7/8 based on current state Master Grid Schedule.

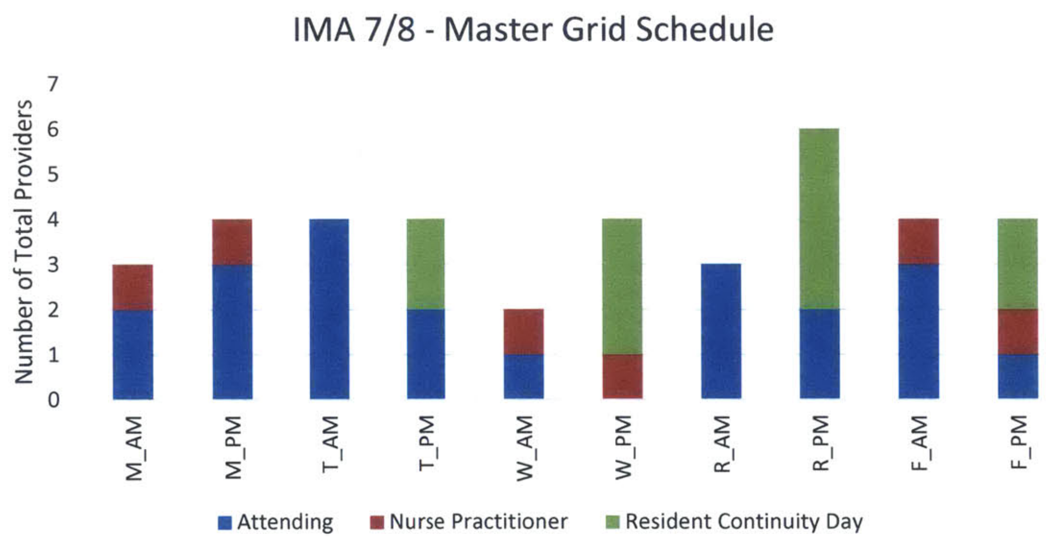


Figure 42 Current state Master Grid Schedule for IMA 7/8.

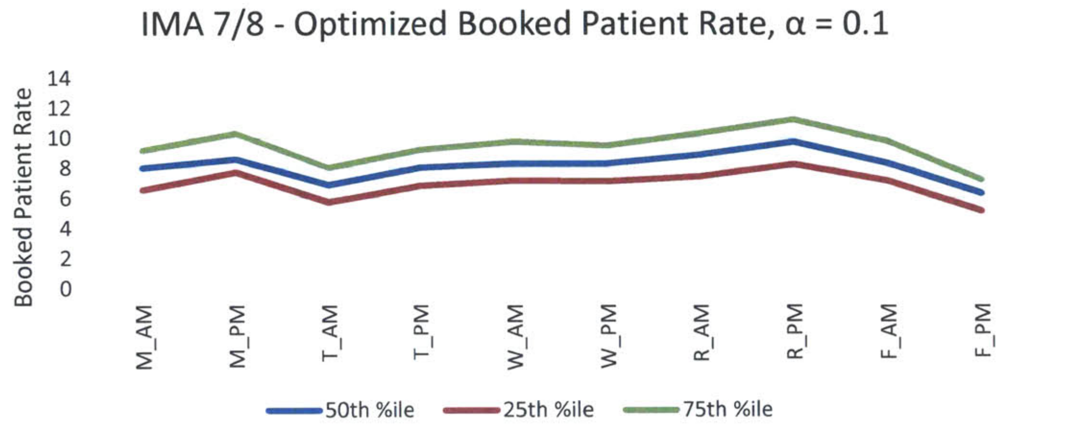


Figure 43 Expected booked patient rate for IMA 7/8 based on optimized Master Grid Schedule.

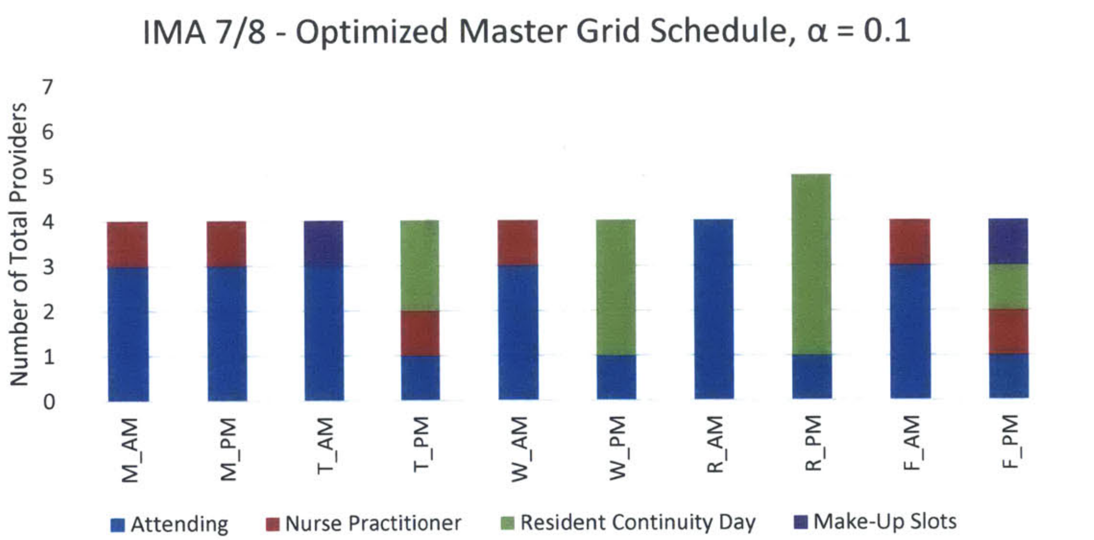


Figure 44 Optimized Master Grid Schedule for IMA 7/8.

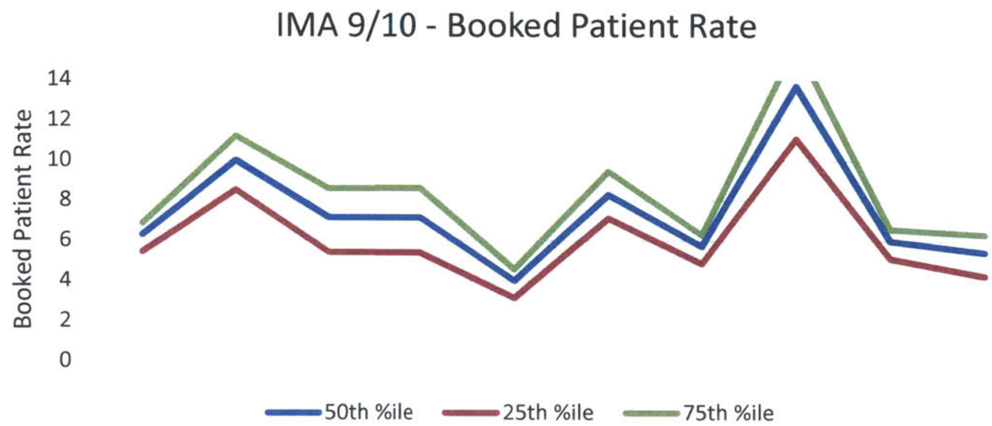


Figure 45 Expected booked patient rate for IMA 9/10 based on current state Master Grid Schedule.

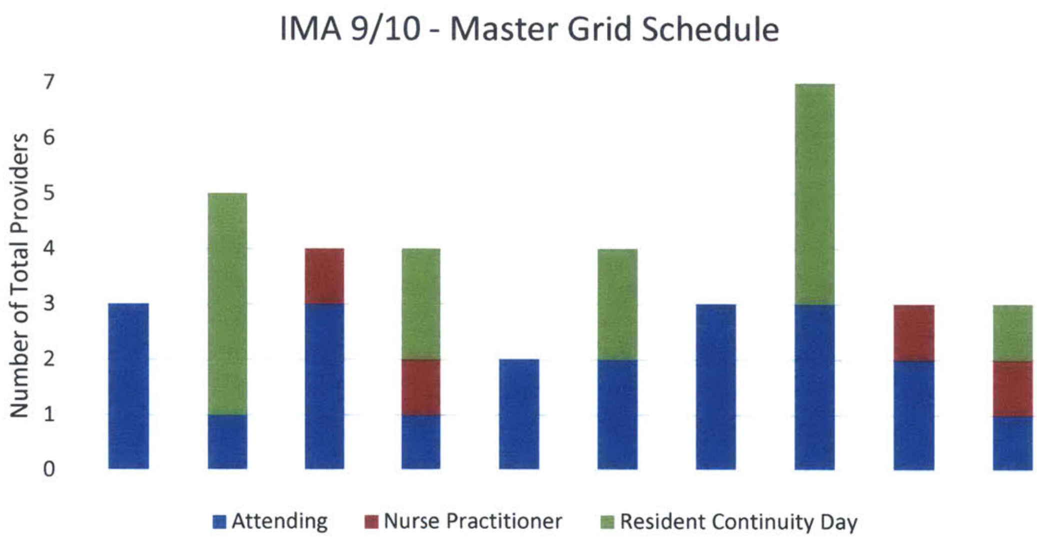


Figure 46 Current state Master Grid Schedule for IMA 9/10.

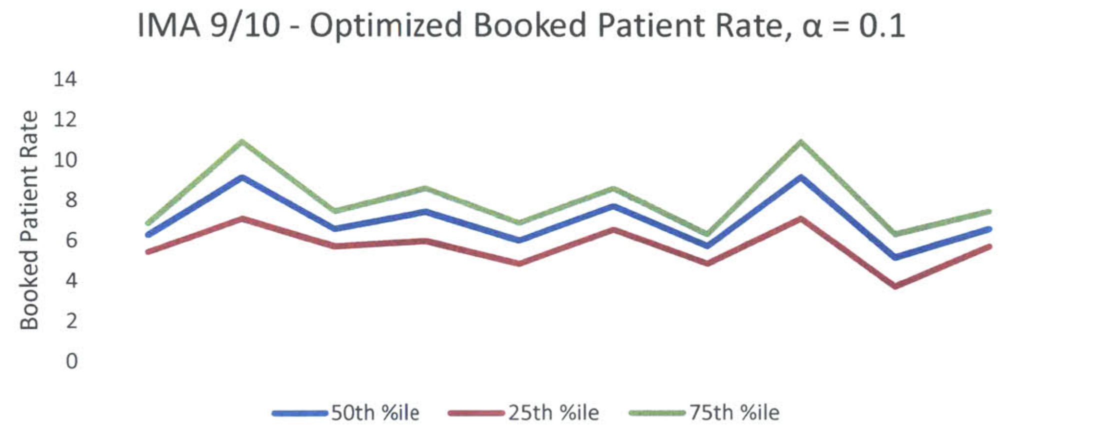


Figure 47 Expected booked patient rate for IMA 9/10 based on optimized Master Grid Schedule.

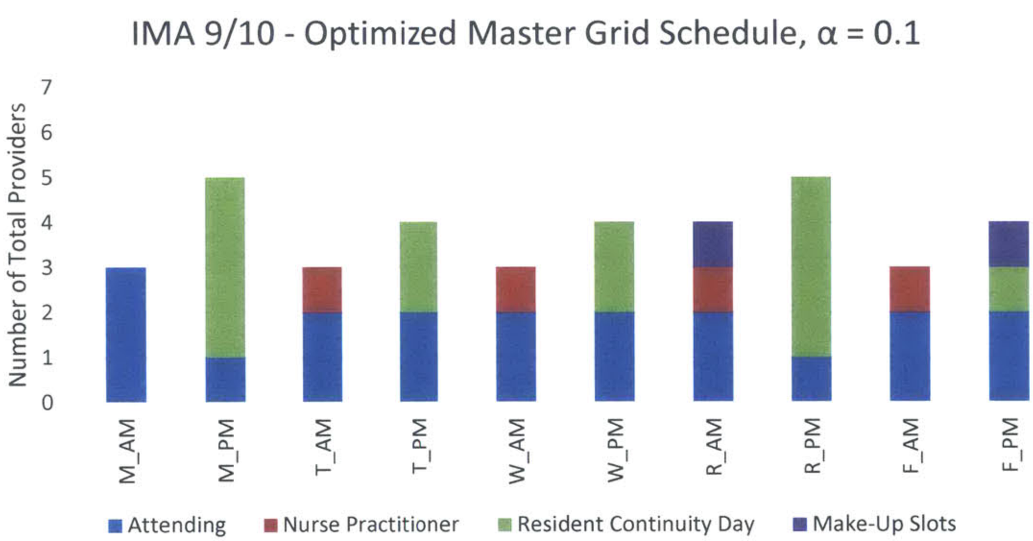


Figure 48 Optimized Master Grid Schedule for IMA 9/10.

Appendix B

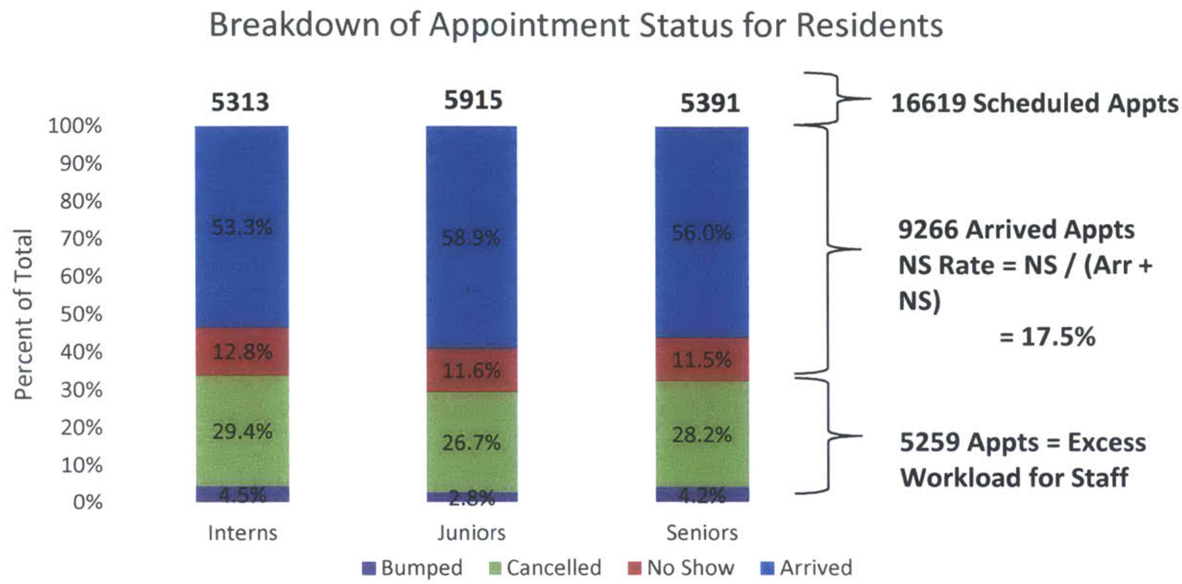


Figure 49 Breakdown of appointment status for residents during June 2012 - June 13.

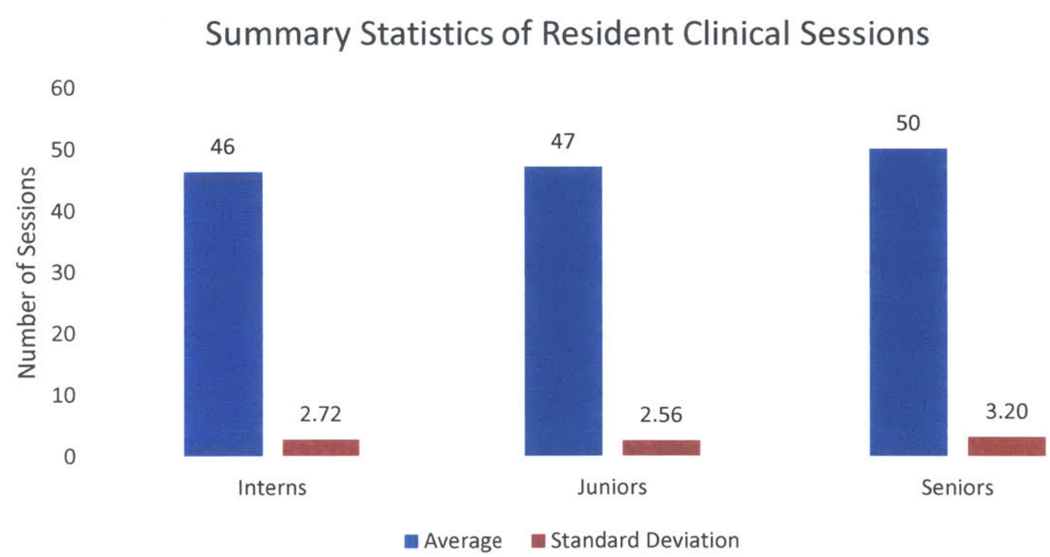


Figure 50 Summary statistics on the number of resident clinical sessions from June 2012 - June 2013.

Bibliography

1. Hospital Overview - Massachusetts General Hospital, Boston, MA. http://www.massgeneral.org/about/overview.aspx. Accessed April 26, 2015.

2. Primary Care -- AAFP Policies. http://www.aafp.org/about/policies/all/primary-care.html. Accessed April 26, 2015.

3. Nijmeijer KJ, Huijsman R, Fabbricotti IN. Creating advantages through franchising in healthcare: a qualitative, multiple embedded case study on the role of the business format. *BMC Health Serv Res.* 2014;14(1). doi:10.1186/s12913-014-0485-5.

4. Bodenheimer T, Wagner EH. Improving Primary Care for Patients With Chronic Illness. 2014;288(14):1775-1779.

5. Grumbach K, Bodenheimer T. Can health care teams improve primary care practice? *JAMA.* 2004;291(10):1246-1251. doi:10.1001/jama.291.10.1246.

6. Shipman SA, Sinsky CA. Expanding primary care capacity by reducing waste and improving the efficiency of care. *Health Aff (Millwood).* 2013;32(11):1990-1997. doi:10.1377/hlthaff.2013.0539.

7. Sinsky C a., Willard-Grace R, Schutzbank AM, Sinsky T a., Margolius D, Bodenheimer T. In search of joy in practice: A report of 23 high-functioning primary care practices. *Ann Fam Med.* 2013;11(3):272-278. doi:10.1370/afm.1531.

8. Nayor MD, Coburn KD, Kurtzman ET et al. *Inter-Professional Team-Based Primary Care for Chronically Ill Adults: State of the Science. White Paper Presented at: The ABIM Foundation Meeting to Advance Team-Based Care for the Chronically Ill in Ambulatory Settings.* Philadelphia, PA; 2010.

9. Mundt MP, Gilchrist VJ, Fleming MF, Zakletskaia LI, Tuan W-J, Beasley JW. Effects of Primary Care Team Social Networks on Quality of Care and Costs for Patients With Cardiovascular Disease. *Ann Fam Med.* 2015;13(2):139-148. doi:10.1370/afm.1754.INTRODUCTION.

10. Zawora MQ, Medical SK, Jefferson T. Turning team-based care into a winning proposition. 2015;64(3):159-165.

11. Taché S, Chapman S. The Expanding Roles and Occupational Characteristics of Medical Assistants: Overview of an Emerging Field in Allied Health. *J Allied Health.* 2006;35(4):233. https://www.questia.com/library/journal/1P3-1191050181/the-expanding-roles-and-occupational-characteristics. Accessed April 26, 2015.

12. Taché S, Hill-Sakurai L. Medical assistants: the invisible "glue" of primary health care practices in the United States? *J Health Organ Manag.* 2010;24(3):288-305. doi:10.1108/14777261011054626.

13. Elder NC, Jacobson CJ, Bolon SK, et al. Patterns of Relating Between Physicians and Medical Assistants in Small Family Medicine Offices. 2014:150-158. doi:10.1370/afm.1581.INTRODUCTION.

14. Defining the PCMH | pcmh.ahrq.gov. http://pcmh.ahrq.gov/page/defining-pcmh. Accessed April 26, 2015.

15. Rosenthal TC. The medical home: growing evidence to support a new approach to primary care. *J Am Board Fam Med.* 2008;21(5):427-440. doi:10.3122/jabfm.2008.05.070287.

16. Aiken LH. Hospital Nurse Staffing and Patient Mortality, Nurse Burnout, and Job Dissatisfaction. *JAMA J Am Med Assoc.* 2002;288(16):1987-1993. doi:10.1001/jama.288.16.1987.

17. Lang TA, Hodge M, Olson V, Romano PS, Kravitz RL. Nurse-patient ratios: a systematic review on the effects of nurse staffing on patient, nurse employee, and hospital outcomes. *J Nurs Adm.* 34(7-8):326-337. http://www.ncbi.nlm.nih.gov/pubmed/15303051. Accessed April 26, 2015.

18. Dickson EW, Singh S, Cheung DS, Wyatt CC, Nugent AS. Application of lean manufacturing techniques in the Emergency Department. *J Emerg Med.* 2009;37(2):177-182. doi:10.1016/j.jemermed.2007.11.108.

19. Jimmerson C, Weber D, Sobek DK. Reducing waste and errors: piloting lean principles at Intermountain Healthcare. *Jt Comm J Qual Patient Saf.* 2005;31(5):249-257. http://www.ncbi.nlm.nih.gov/pubmed/15960015. Accessed April 26, 2015.

20. Kibbe DC, Bentz E, McLaughlin CP. Continuous quality improvement for continuity of care. *J Fam Pract.* 1993;36(3):304-308. http://www.ncbi.nlm.nih.gov/pubmed/8454977. Accessed April 26, 2015.

21. Hummer J, Cristina D. *Improvement in Prescription Renewal Handling by Application of the Lean Process.*

22. Grove AL, Meredith JO, Macintyre M, Angelis J, Neailey K. Lean implementation in primary care health visiting services in National Health Service UK. *Qual Saf Health Care.* 2010;19(5):e43.

23. Schäfer WLA, Boerma GW, Murante AM, Sixma JM, Schellevis G. Assessing the potential for improvement of primary care in 34 countries : a cross-sectional survey. *Bull World Health Organ.* 2015;93(3):161-168.

24. Christensen BA. Improving ICU patient flow through discrete-event simulation. 2012.

25. Range AR. Improving surgical patient flow through simulation of scheduling heuristics. 2013.

26. Holt CC, Modigliani F, Simon HA. A Linear Decision Rule for Production and Employment Scheduling. *Manage Sci.* 1955;2(1):1-30. doi:10.1287/mnsc.2.1.1.

27. Fuller E. Tackling scheduling problems using integer programming. 1998.

28. Aickelin U, Dowsland KA. An indirect Genetic Algorithm for a nurse-scheduling problem. *Comput Oper Res.* 2004;31(5):761-778. doi:10.1016/S0305-0548(03)00034-0.

29. Attending Physician – Doctors – Health Care Team – ECHO Resources. http://www.ecfmg.org/echo/team-doctors-attending-physician.html. Accessed April 26, 2015.

30. ACGME > Program and Institutional Accreditation > Medical Specialties > Internal Medicine. https://www.acgme.org/acgmeweb/tabid/134/ProgramandInstitutionalAccreditation/MedicalSpecialties/InternalMedicine.aspx. Accessed April 26, 2015.

31. How to Become a Nurse. http://www.nursingworld.org/EspeciallyForYou/What-is-Nursing/Tools-You-Need/RegisteredNurseLicensing.html. Accessed April 26, 2015.

32. AANP - What's an NP? http://www.aanp.org/all-about-nps/what-is-an-np. Accessed April 26, 2015.

33. AAMA - What is a Medical Assistant. http://www.aama-ntl.org/medical-assisting/what-is-a-medical-assistant#.VTOynSFVikp. Accessed April 26, 2015.

34. Mitchell P, Wynia M, Golden R, et al. *Core Principles and Values of Effective Team-Based Health Care.* Washington D.C.; 2012. www.iom.edu/tbc.

35. ACGME > About > Newsroom > Fact Sheet. https://www.acgme.org/acgmeweb/About/Newsroom/FactSheet.aspx. Accessed April 26, 2015.