Pooling and Segmentation to Improve Primary Care Prescription Management

by

Thomas Daniel Sanderson

Submitted to the MIT Sloan School of Management and the Engineering Systems Division on May 9, 2014, in partial fulfillment of the requirements for the degrees of Master of Business Administration and Master of Science in Engineering Systems Division

Abstract

Analyses of schedule history and medical records for large primary care medical practice are combined with time studies to develop a quantitative network flow model of the prescription management process, including metrics for the yield of prescription requests, excess requests for chronic & stable medications, and prescriptions that overflow from scheduled appointments. The model is used to estimate the impact of segmenting and pooling the prescription workflows into a central prescription management group, resulting in recommendations for a new prescription management system. Interventions on pharmacy/practice coordination for faxes, modification of the voicemail system, and improved workflow are piloted to validate model estimates. The recommended changes are expected to improve prescription response time and accuracy, reduce resource utilization, improve patient medicine compliance and, ultimately, patient health outcomes.

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

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

This work would not have been possible without the support of many contributors.

First and foremost, my eternal thanks to my wife, Carson, and my children: Daniel, Riley, Audrey, and Cara. Without their support and understanding, I could not have participated in the Leaders For Global Operations (LGO) program or performed this work at MGH.

Prof. Retsef Levi’s dedication to better healthcare operations and always keeping the patient at the forefront was personally inspiring and professionally motivating. His tireless work, including supporting this project on a near-daily basis late into the night by teleconference from Israel, made the MIT/MGH collaboration, our work with Internal Medicine Associates, and this project, possible.

Prof. David Simchi-Levi provided valuable support from the beginning to keep the goal of this completed thesis in sight, to keep the right balance of business and technical rigour, and to work seamlessly with Prof. Levi to navigate the complex waters of the LGO dual-degree requirements.

Dr. Peter Dunn & Bethany Daily of MGH Perioperative Services placed their faith in MIT and the LGO program to become partners, expanded beyond their own group for the selfless benefit of MGH and all patients, expressed their confidence in an eager but unfamiliar aerospace engineer, and provided a welcome voice of experience and hospitality to help navigate the culture shock of the healthcare landscape.

Brianna Germain (Perioperative Services), Carine Yelibi (Primary Care Operational Improvement), and Mary Morgan (Lab of Computer Science) provided seamless access and responded to unending questions regarding that holy grail of any operations project: administration and data!

Virginia Manzella, Dr. Sherry Haydock, Dr. Blair Fosburgh, and Dr. Daniel Horn of Internal Medicine Associates opened their arms and offices to our team, willing to ask hard questions and receive hard answers, and always put the good of their patients at the forefront to allow outsiders to come in to help make better what they have dedicated so much of their lives to.

Kaitlyn Kelley, Yolanda Fernandez, and Morgan Lichtenstein of Internal Medicine As-
sociates eagerly showed us the ropes and provided invaluable honest feedback from the
front-lines, at the expense of their own free time, when theory and reality collided.

Prof. Kate Kellogg of the Sloan School of Management provided valuable insight into the
“OP rich environment” of IMA.

Last, but by no means least, Cecilia Zenteno of the Sloan School of Management pro-
vided expertise, guidance, boundaries, support, friendship, discipline, a welcome perspective,
mastery of R and data processing, and a nearly unlimited supply of good coffee.
## Contents

1 Introduction 15  
1.1 Background ......................................................... 15  
1.1.1 Massachusetts General Hospital .............................. 15  
1.1.2 Primary Care .................................................... 16  
1.1.3 Internal Medicine Associates ................................. 17  
1.2 Project Overview .................................................. 17  

2 Literature Review 23  
2.1 The Role Of Primary Care ........................................ 23  
2.2 Primary Care Prescription Management .......................... 24  
2.2.1 Drug Adherence .................................................. 24  
2.2.2 Renewal Errors .................................................. 27  
2.3 Improvement Due To Process Redesign ........................... 27  
2.3.1 Use Of Simulation ............................................... 28  
2.3.2 Interruption ...................................................... 28  
2.3.3 Centralization ................................................... 29  
2.3.4 Specialization ................................................... 30  
2.4 Proactive Medication Management .................................. 30  

3 Characterization of Resource Use In IMA 33  
3.1 Introduction ......................................................... 33  
3.2 Time Study ........................................................ 34  
3.3 Methods .......................................................... 37
A  Role Observation Form  83
B  Patient Observation Form  87
C  Revised Renewal Line Voicemail Script  91
D  Additional Time Study Observations  93
E  Network Flow Model Screenshots  97
List of Figures

1-1 IMA Organizational Structure – each pod consists of a small independent group of attending doctors and their related support staff .......................... 18
1-2 MGH information system data sources & data processing relationships. Source – MIT direct observation. ................................................................. 19
1-3 Prescription system drug adherence feedback loops. Note that all “No” arcs are generally invisible to the practice or doctor such that non-adherence is only identifiable when the patient fails to renew on time. .............................. 21
3-1 Fraction of observed time spent managing prescriptions. Source – MIT direct observations. ................................................................. 36
3-2 IMA schedule performance. Delay – difference between scheduled and actual start time. Overrun – difference between scheduled and actual duration. Each marker corresponds to a single appointment, with like symbols corresponding to the same provider. ................................................................. 41
3-3 Time distribution by task – Patient Service Coordinator (PSC) .......................... 44
3-4 Time distribution by task – Medical Assistant (MA) ........................................ 44
3-5 Time distribution by task – Nurse (RN) .......................................................... 45
3-6 Time distribution by task – Nurse Practitioner (NP) ....................................... 45
3-7 Time distribution by task – Resident ............................................................... 46
3-8 Time distribution by task – Attending (MD) .................................................... 46
3-9 Time distribution by task – Preceptor ............................................................. 47
4-1 IMA current state prescription workflow - single pod ........................................ 50
4-2 IMA current state prescription workflow - all pods ........................................... 50
List of Tables

3.1 Total role observations ........................................... 40
4.1 Requested vs. issued prescriptions - normalized to 2012 total request volume 56
5.1 Future state metrics ................................................. 79
Chapter 1

Introduction

Like many other healthcare delivery systems, Massachusetts General Hospital (MGH) anticipates that its primary care practices will play a major role in slowing the growth of health care costs through optimal management of distinct patient populations. Internal Medicine Associates (IMA) is the largest primary care practice at MGH. As the first MIT Leaders for Global Operations (LGO) engagement with IMA, this project had three primary goals: (i) quantify how IMA currently utilizes resources; (ii) identify high-leverage areas for system redesign; and (iii) develop a concrete implementation plan for one area.

1.1 Background

1.1.1 Massachusetts General Hospital

Massachusetts General Hospital (MGH) is the largest and oldest hospital in New England, and is consistently ranked as one of the best hospitals in the country [1]. MGH handles approximately 48,000 inpatients and 1.5 million outpatients annually, including: 90,000 ER visits, 37,000 surgeries, and 3,600 deliveries. MGH originally partnered with the LGO program to improve operations in the Perioperative Services group, which supports all surgical specialties. Based on several successful prior projects ([14][13][39][38][42]), in 2013 the MGH/MIT collaboration extended to include MGH Primary Care.
1.1.2 Primary Care

Primary care practices provide long-term longitudinal care for patients, coordinating multiple specialists and referrals as necessary, and provide an important training environment for medical residents. Historically, primary care providers have been the "gateway" to specialists with separate diagnoses/billing/treatment/follow-up procedures for each group. However, rising medical costs and increasingly complex treatment options have highlighted the need for better coordination among various caregivers for a single patient. As a result, insurance companies are turning to primary care providers to integrate all of the various specialists and treatments in the hopes of providing more cost-effective care, and modifying their reimbursement policies to match. Accountable Care Organizations (ACOs) are groups of doctors, hospitals, and other healthcare providers who come together under a unified Medicare reimbursement framework where, rather than fee-for-service to each provider, the entire organization is compensated based on overall health outcomes. An analogous non-Medicare structure, capitation, reimburses providers on a risk-adjusted per capita basis regardless of the number or type of treatments required for any given patient. Both ACOs and capitated plans use primary care as the lynch-pin of coordination and integration between the vast array of other health care providers that a patient may require. This shift will change the operating framework of primary care away from individual patient encounters and towards managing health outcomes. The transition of primary care to a Patient-Centered Medical Home (PCMH) model to support outcome-based reimbursement plans places primary care at the forefront of healthcare cost containment efforts.

MGH provides comprehensive primary care services for 200,000 patients at 15 locations [2]. MGH is also the original and largest teaching hospital of Harvard Medical School [1]. Like many other primary care practices [46], MGH has observed increasing dissatisfaction and burnout among primary care physicians due to stressful and chaotic work environments, increasing administrative and regulatory demands, fragmentation of care delivery, and greater than ever expectations placed on primary care providers. MGH is redesigning their primary care system to increase capacity, improve utilization, and resolve the sustainability of primary care. One MGH primary care practice, Internal Medicine Associates, requested...
MIT’s assistance with this redesign due to the scope and scale of the practice and their acute resource utilization challenges.

1.1.3 Internal Medicine Associates

Internal Medicine Associates (IMA), the largest and oldest primary care practice at MGH, cares for 31,000 patients with a staff of 46 attending physicians, 59 resident physicians, 6 nurse practitioners, and 86 related clinical and administrative support staff.

IMA was originally created to provide a training environment to complete the primary care residency requirement for students at Harvard Medical School. Over time, the practice has grown as both the training and patient care needs of the MGH population increased. Throughout the years, groups of attending physicians have split off from IMA to form other primary care practices within MGH.

IMA is characterized by a unique parallel operations structure with 14 semi-independent “pods”, shown schematically in Figure 1-1. This parallel structure has made IMA highly scalable but presents significant challenges in performance tracking and change management due to the autonomous nature of each pod. IMA’s scale suggests that typical operations management techniques, such as segmentation and pooling, may be useful to improve performance and reduce waste.

As the largest, oldest, and most complex primary care practice at MGH, IMA is both the flagship for primary care innovation but also the most difficult implementation environment. MGH has expressed a desire to increase the capacity (patients per IMA headcount) and productivity (revenue per provider) at IMA to accommodate the increasing workload in the changing primary care environment discussed in section 1.1.2. Due to its scale, success in operational improvement at IMA will make a significant difference in patient care, and viability of primary care, at MGH.

1.2 Project Overview

The first phase of this project (Section 3.2) quantifies how IMA currently utilizes resources. This phase required extensive data analysis, leveraging existing data sources through the
Figure 1-1: IMA Organizational Structure – each pod consists of a small independent group of attending doctors and their related support staff.
Figure 1-2: MGH information system data sources & data processing relationships. Source — MIT direct observation.

MGH Primary Care Office Insight (PCOI) system and OnCall electronic medical record. These systems provide detailed logs of all patients, appointments, medications, and prescriptions from 2010-2012. Figure 1-2 shows the relationship between different data sources, some of which were created as part of this work. In addition, 127 hours of detailed quantitative time studies were performed with all staff roles across IMA to determine what aspects of the current primary care model are consuming disproportionate practice resources.

The second phase of the project (Section 3.4) uses the results of the time studies & data analysis to identify the highest leverage process changes for potential system redesign efforts. Five high-leverage areas were proposed to the practice to select the first area for implementation with an assessment of impact, feasibility, and interaction with other practice
operations. This phase identifies prescription management as the first focus area due to: significant resource utilization across all roles; widespread dissatisfaction with the current state; lack of coupling with the compensation and scheduling processes; and relatively high process control within IMA. The latter two points, in particular, simplify implementation. The remaining focus areas (rescheduling load, scheduled appointment duration, patient panel capacity analysis, and support staff capacity analysis) are discussed in more detail in Section 5.5.

The third phase (Chapter 4) uses the results from the prior phases, plus additional direct observation of the current drug prescription management process, to develop a network flow model of the current state in the clinic. Model validation was determined by the match between predicted and observed resource utilization for a given prescription load. Once calibrated, the model was used to quantify the improvement based on various centralization options, developed in collaboration with the practice, for pooling and segmentation of the various prescription management workflows. These quantified improvements were used to select and develop the implementation plan for a new prescription management system. The model was also used to study the opportunities for implementing proactive drug management to reduce prescription demand variance and volume, and to improve patient medical adherence.

Due to the lack of system feedback for patient actions in the current prescription process (Figure 1-3), we had to focus on operational/efficiency metrics derived from the interaction of schedule, prescription, and drug data. Historical logs of prescription output are combined with three newly developed metrics: redundant/non-actionable prescription requests (Yield), prescription requests done outside an appointment that could have been done at an appointment (Spillover), and prematurely renewed prescriptions for chronic medications (Excess chronic & stable). A network flow model estimates the resources required to process non-appointment prescriptions under a variety of scenarios, measured in full-time-equivalent (FTE) staff. One FTE is one person working forty hours per week, fifty weeks per year (two weeks vacation per year are assumed).

Modeling shows that IMA currently expends 4.1 FTE to service non-appointment prescriptions. This work is spread out part-time over an average of 28 nurses and medical
Figure 1-3: Prescription system drug adherence feedback loops. Note that all "No" arcs are generally invisible to the practice or doctor such that non-adherence is only identifiable when the patient fails to renew on time.
assistants, with work arriving at random intervals throughout the day and disrupting other tasks. The 4.1 FTE estimate is a lower bound as it does not include the additional cost of these disruptions. 15% of prescription requests are redundant/non-actionable, 4% of prescription requests could be done in an appointment but are not, and 8% of prescription requests are due to chronic prescriptions being renewed early. A new prescription system is developed to segment prescriptions away from other processes and centralize the prescription streams of multiple pods into one specialized group. Based on the expected reduction in variance due to pooling, improved cycle times due to specialization, and the opportunity to proactively identify and correct non-value-added prescription requests, we predict that 2–2.5 FTE will be capable of handling the entire non-appointment prescription load, a 37% reduction in resource utilization that will free 28 nurses and medical assistants from frequent interruption to allow them to focus on direct patient care.

A pilot intervention with a major New England retail pharmacy chain showed that proper prescription linking could reduce redundant/non-actionable prescriptions requests by 35 to 50% (P-value < 0.05). The voicemail renewal lines will be modified to improve yield, then eventually eliminated. Two dedicated staff will be added to the practice in 2014 to implement a centralized prescription management group, forming the foundation for proactive monitoring and management of patient medicine compliance and drug-related health outcomes.
Chapter 2

Literature Review

2.1 The Role Of Primary Care

Primary care provides, directs, and arranges for all of a patient’s healthcare needs [2]. Recent perspectives of primary care include the intention that preventative care coupled with long-term management of chronic conditions will reduce the later use of high-cost high-risk medical interventions. In particular, there is extensive literature that appropriate primary care can improve outcomes and lower costs for many common chronic conditions: hypertension [10][29], diabetes [9], and cholesterol [29][12]. The quantitative savings of primary care on many other comorbidities is an area of active research.

At the same time, primary care is facing a looming shortage of doctors. Relative to the general population, primary care physicians are more likely to have symptoms of burnout and be dissatisfied with worklife [44]. Most solutions proposed for the looming shortage of primary care physicians entail strategies that fall into one of three categories: (i) train more; (ii) lose fewer; or (iii) find someone else. A fourth strategy is now receiving more attention based on inefficiency and waste in primary care today: improve productivity of the available doctors to allow them to keep pace with the workload. If widely adopted, small efforts to empower non-physicians, re-engineer workflows, exploit technology, and update policies to eliminate wasted effort could yield the capacity for millions of additional patient visits per year in the United States [45].

Practices that do show high performance and improved physician satisfaction show con-
sistent characteristics: proactive planned care, with previsit planning and previsit laboratory
tests; sharing clinical care among a team, with expanded rooming protocols, standing orders,
and panel management; sharing clerical tasks with collaborative documentation (scribing),
nonphysician order entry, and streamlined prescription management; and improving team
functioning through co-location, team meetings, and work flow mapping [46].

In the context of this work, any improvement in primary care operations provides an
opportunity to reduce resource utilization, at least on a per patient basis. This may create
the opportunity to apply that freed capacity to improve care and/or mitigate the primary
care doctor shortage.

2.2 Primary Care Prescription Management

Within the primary care environment, the issue and maintenance of prescriptions consumes
a significant amount of resources and impacts the majority of patients (Chapter 4) but can
have major positive impacts on patient outcomes. For example, national survey data suggests
that statin therapy resulted in roughly 40,000 fewer deaths, 60,000 fewer hospitalizations
for heart attacks, and 22,000 fewer hospitalizations for strokes in 2008 [21]. As a result,
drug adherence (the degree to which patients comply with the intended drug therapy) and
prescriptions errors are of central concern to any primary care prescription management
system. Operational improvements discussed in this work can free resources to improve the
quality and responsiveness of the prescription system, improving health outcomes through
better adherence and decreasing costs by serving the existing demand more efficiently.

2.2.1 Drug Adherence

Adherence to the medical regimen continues to rank as a major clinical problem in the
management of patients treated with drugs and life-style modification [33]. There is potential
for increased drug utilization to provide a net economic return when it is driven by improved
adherence with guidelines-based therapy [48]. Medication nonadherence is a growing concern
to clinicians, healthcare systems, and other stakeholders (e.g. payers) because of mounting
evidence that it is prevalent and associated with adverse outcomes and higher costs of care
[24]. Helping patients take their existing medication as prescribed is identified as one of the most effective medical interventions to significantly improve the health of patients although it does not require the latest technology or expensive medication [11].

One study showed that 66% of all Americans fail to take any of their prescription medicines [22]. A pharmacy care program that coordinated primary care doctors, pharmacists, and patients led to increases in medication adherence, medication persistence, and clinically meaningful reductions in blood pressure, whereas discontinuation of the program was associated with decreased medication adherence and persistence [29]. A similar study found dyslipidaemia is a poorly-controlled condition in clinical practice largely because more than 10% of patients are potentially non-adherent [56]. Chronic obstructive pulmonary disease is frequently treated in a primary care setting; one study found 48% of patients underadherent and 5% overadherent and found that age and number of drugs increase the chance of underadherence [30]. Another study of antihypertensive prescriptions found 31.8% non-adherence [57]. Among patients discharged with prescriptions for aspirin, statin, and blockers after acute myocardial infarction, 34% of patients stopped at least 1 medication and 12% stopped all 3 medications within 1 month of hospital discharge [24]. Another study demonstrated that only 40% of patients were still taking statin medications 2 years after hospitalization for acute coronary syndrome and adherence was even lower for patients taking statins for chronic coronary artery disease [24]. An antibiotic study found 20.7% of patients failed to take the complete course of antibiotic treatment [40].

Simply providing patients with medication timetables and computer generated consumer product information does not improve drug adherence in primary care [8]. In fact, even filling prescriptions is a challenge. One study of a large database of electronic prescriptions found only 78% of all prescriptions were filled. Of e-prescriptions for new medications, only 72% were filled. Efforts to increase adherence could dramatically improve the effectiveness of medication therapy [20]. In 2004, an estimated 1.6 million Medicare beneficiaries (4.4%) failed to fill or refill 1 or more prescriptions [28].

Metrics do exist to attempt to identify drug abuse, such as the Doctor Shopping Indicator (DSI) used to assess the relative abuse liability of benzodiazepines and to detect signals of new patterns of abuse in settings where centralized records of prescription or deliveries
are available for the great majority of patients [37]. Other studies suggest that analysis of prescription logs can assess medical compliance and correlates with patient drug effects in general [50]; and that younger subjects, new patients, and those without comorbidities should receive more meticulous monitoring of their medication-taking behavior [56]. Traditional methods of assessing non-adherence, mainly self-reporting, may be ineffective as agreement between self-reported and refill adherences is poor to fair [34]. Currently, the two most commonly used measures of medication adherence based on pharmacy data are the medication possession ratio and the proportion of days covered methods, which essentially are defined by the number of doses dispensed in relation to a dispensing period [24]. These are similar to, but less specific, than the “excess chronic & stable” metric developed in Section 4.3.4.

Improper and unnecessary use of medications cost over $200 billion in the US in 2012: patients fail to receive the right medications at the right time or in the right way, or receive them but fail to take them. Improvements are possible only through collaboration among multiple healthcare stakeholders: providers, pharmacists, patients, payers, pharmaceutical manufacturers and policymakers. In addition, healthcare informatics — the use of technology and analytical approaches to harness the value of data — is a key driver of improvements [5].

Some interventions have proven to be at least partially effective in addressing this issue. A pharmacy care program led to increases in medication adherence, medication persistence, and clinically meaningful reductions in blood pressure, whereas discontinuation of the program was associated with decreased medication adherence and persistence [29]. However, one study found that pharmacist contact with the patient by phone or fax was no better than usual care [35], suggesting that the nature of pharmacist contact is important.

Plainly, improper prescribing and drug adherence imposes a significant burden on the healthcare system in terms of poorer outcomes and higher costs. The operational improvements and new metrics discussed in this work provide a framework to improve the accuracy and timeliness of prescriptions, and to detect potential non-adherence earlier and more effectively.
2.2.2 Renewal Errors

Some drug adherence issues are the result of incorrect or overly complex prescription management systems. Poorer-quality prescribing and primary care are associated with higher medical spending in general and inpatient spending in particular [60]. Because many patients have multiple chronic conditions, therapeutic regimens often involve multiple medications and frequent daily dosing. Such regimen complexity may undermine effective chronic disease management. Patients who are prescribed medications that must be taken multiple times per day are less likely to adhere to their treatments than patients with simpler dosing schedules. Interventions that simplify treatment regimens by reducing dosing frequency or by switching patients to fixed-dose medication combinations result in substantial improvements in appropriate medication use [13]. Greater prescribing and filling complexity is associated with lower levels of adherence. Patients with the least refill consolidation had adherence rates that were 8% lower over the subsequent year than patients with the greatest refill consolidation [13]. A large study of safety reports found medication incident reports represented 9.68% of all patient safety incidents. There were relatively smaller numbers of medication incident reports from primary care, representing 8.5% of the total [15]. A UK study that evaluates the efficacy of pharmacist-led, information technology-based intervention for medication errors found it is an effective method in lessening the prescription errors [6].

2.3 Improvement Due To Process Redesign

There is considerable literature on the use of operations management principles (chiefly modeling and process redesign) in the healthcare field, although application to the outpatient primary care environment is significantly less common. One case of the application of lean principles to an outpatient pharmacy also showed many recommendations similar to those discussed in this work [26]:

- Answer more calls “live” whenever possible, rather than taking messages.
- Relocate the pharmacy technician and the MA into the same office space
- Utilize clerical staff to move necessary documents between the clinic and the renewal office.

- Change the phone tree message for patients with clearer directions

This work, in part, explores the use and applicability of these techniques to the primary care practice prescription management environment and extends them via modeling to provide quantitative predictions of expected performance after redesign.

2.3.1 Use Of Simulation

Discrete event simulation is a common technique to analyze the current state of healthcare operations and suggest improvements. At MGH, a SICU and its six primary downstream units were modeled in a highly detailed discrete event simulation. The validated simulation is applied to evaluate the impact of several possible operational adjustments [14]. Another study decreased the patient processing time and length of stay in the PACU through the application of discrete event simulation [43]. Other simulation models are built to evaluate the downstream effects of scheduling rules and discharge process changes in the perioperative environment [39][38].

Two known cases of simulation in an ambulatory care clinic, which is roughly comparable to the primary care environment, also used discrete event simulation [49][41].

2.3.2 Interruption

As discussed in detail in Chapter 3, primary care operations at IMA are characterized by frequent interruption of tasks, especially for the support staff, by many different workflows throughout the day. Interruptions have been described as extremely disruptive, as “constant, constant, multitasking craziness”, and as events that can cost a knowledge worker as much as 15 minutes in recovery time [31]. In one study, subjects’ responses are substantially slower and, usually, more error-prone immediately after a task switch [32]. In busy interrupt-driven clinical environments, clinicians reduce the time they spend on clinical tasks if they experience interruptions, and may delay or fail to return to a significant portion of interrupted
tasks. Task shortening may occur because interrupted tasks are truncated to “catch up” for lost time, which may have significant implications for patient safety [55].

2.3.3 Centralization

The new prescription management system proposed in Chapter 4 relies, in part, on taking advantage of centralization to enable pooling benefits. In this context, pooling refers to combining two or more independent but similar workstreams into one stream handled by one entity. This provides two primary benefits: (i) the variability, measured by standard deviation or coefficient of variation, of the combined workload is lower than the sum of the individual workloads; and (ii) the combined workload is larger, allowing the central processing entity to devote a greater fraction of their capacity to the pooled workstream. There is considerable research on the positive benefits of centralization in the healthcare environment.

Uniform standards, vocabularies, and centralized knowledge structures and services could reduce rework by vendors and care providers, improve dissemination of well-constructed interventions, and accelerate the movement of new medical knowledge from research to practice [54]. Centralization of coronary revascularization procedures in two Calgary hospitals resulted in a higher rate of procedures within shorter hospital stays and more-favorable short-term outcomes despite sicker patients [23]. Centralization of esophageal cancer surgery improved outcomes even with more severe comorbidity [58]. Specialist cancer centers that centralized surgical care reduced the number of adverse events in patients with pancreatic cancer [59]. Centralization of pharmacy services results in decreased inventory costs [53]. A centralized prescription network with filled prescription feedback reduced inappropriate prescription filling of opioids by 32.8% and benzodiazepines by 48.6% [19]. Centralized pharmaceutical services, pharmacist consultation and drug use review significantly reduced drug use and medication costs [51]. Centralization of travel medicine service reduced pharmacy costs by 17% due to better adherence to expert recommendations [7].
2.3.4 Specialization

In addition to pooling benefits, centralization allows employees to specialize in their tasks. One primary-care redesign that included on-site physician pharmacy dispensing improved patient medication adherence [52]. Application of operations management principles to improve quality through specialized employees and processes increased accountability, drove improvement, and produced savings [27]. Care guides, a specialized non-licensed care role, can improve care for some patients with chronic disease at low cost [3][4]. Drug experts, such as pharmacist or pharmacy technicians, can improve prescribing, reduce health-care utilization and medication costs, and contribute to clinical improvements in many chronic medical conditions, such as cardiovascular disease, diabetes, and psychiatric illness [18].

By centralizing prescription management in the primary care environment, this work seeks to gain the benefits of less interruption and greater specialization that have been observed in other areas of health care operations.

2.4 Proactive Medication Management

A future benefit of creating a centralized prescription management group is the capability to switch from reactive to proactive medication management. The use of electronic medical record data to generate information regarding control status and successful drug regimens used by similar patients to develop collective experience decision support is both technically possible and promising [16]. Thoughtful implementation and strategic workflow redesign can mitigate the disproportionate electronic health record related work burden for clinicians, as well as facilitate population-based care [25]. One study of a pharmacist intervention to proactively improve patient health literacy improved adherence in patients with adequate and inadequate health literacy [36]. In a large health maintenance organization, a case-management system was considerably more effective than usual medical care for modification of coronary risk factors after myocardial infarction [17].

Medication synchronization also becomes possible with a central prescription group. Synchronized, bundled prescription renewal is a systematic approach to prescription management that can decrease patient inconvenience, support medication adherence, and save one to two
hours of physician and staff time daily [47].
Chapter 3

Characterization of Resource Use In IMA

3.1 Introduction

Internal Medicine Associates (IMA) is characterized by many semi-autonomous pods. Each pod has its own team of doctors and staff who serve non-overlapping panels of patients. Based on interviews with key IMA members, we suspect considerable variation between and inside pods with respect to roles and responsibilities of different staff members and task flows. The organizational structure of IMA was previously illustrated in Figure 1-1. Each pod typically contains the following roles:

**Attending (MD):** A primary care doctor who has completed their residency and has a panel of patients that they care for in the practice. MDs may be part-time or full-time depending on the number of clinical sessions they have per week, and may have other duties within the hospital or surrounding research environments. MDs diagnose and treat patients, and provide guidance and medical orders to support staff.

**Preceptor:** An attending who oversees the medical actions of one or more residents.

**Resident:** A doctor who has completed medical school but is in some phase of their three year residency. All medical residents perform at least some primary care work, although
many will eventually assume other specialties. All residents at IMA have a small patient panel and perform essentially the same medical tasks as an attending but all of their work must be reviewed and approved by a Preceptor.

**Nurse Practitioner (NP):** An advanced-degree nurse with some diagnostic and prescription authority who can provide autonomous healthcare decisions for less complex medical issues but may require consultation from an MD for more complex issues. Some NPs may also have their own patient panel.

**Nurse (RN):** A licensed medical provider with a nursing degree, trained to perform many medical procedures under the orders of an MD. RNs triage patients as needed, over the phone or on arrival to the practice, and provide intra-visit followups and patient education. RNs do not have prescription authority.

**Medical Assistant (MA):** A non-licensed medical support staff with some training in basic diagnostic procedures, such as taking height/weight/blood pressure/pulse. MAs help prep the patients for appointments, escort them to/from the exam rooms, assist the doctors during the appointment as needed, and generally facilitate the physical flow of patients through the practice. They also perform the administrative processing of prescriptions, referrals, prior authorizations, and related patient paperwork.

**Patient Service Coordinator (PSC):** A non-licensed administrative support staff who handles practice communications (faxes, phone calls, walk-ins); checks patients in and out at the front desks; preps and processes charts and billing forms; and performs all doctor and patient scheduling. They also work together with the MAs to process administrative paperwork for patients.

### 3.2 Time Study

Existing datasets from the PCOI group (Figure 1-2) provide excellent visibility of scheduled patient load and actual patient arrivals, but do not capture work not associated with an appointment, the breakdown of effort required by different roles during appointments, or the difference between scheduled and actual appointment duration.
To fill in this missing information, we conducted a structured time study of 36 members of the practice across seven roles (PSC, MA, RN, NP, resident, MD, and preceptor) across all pods. This effort generated 127 hours of data including task, duration, physical location, and interaction with other members of the practice. This provided a detailed view of the practice current state: workflows, task allocations, staff interaction, and “pain points”. This effort also helped meet the goal of mapping out major challenges and opportunities, and analysis of the resulting data quantitatively confirmed several hypotheses based on observed workflow and interviews with IMA team members:

- Prescription management load falls primarily on the nurses and medical assistants (Figure 3-1)
- Pre- and post-visit work falls primarily on the PSCs and attendings
- Non-visit care forms the majority of the nurse workload

However, the data also shows a significant fraction of time for non-nurse roles spent on non-visit care. This has significant implications for practice reimbursement under a fee-for-service model, as most of this work is not billable, but under future Accountable Care Organization/capitated reimbursement this type of non-visit care will be essential.

The data also shows a statistically significant tendency for appointments to start late, to run over, and a positive correlation between delay and overrun (Figure 3-2). This quantitatively validates the providers’ and patients’ anecdotal experience of waiting room delays and providers’ motive to “give patients a more thorough appointment if they’ve been kept waiting.”

Other general findings include:

- Very low level of non-work activity (breaks, lunch, etc.): staff spend essentially all of their time directly performing IMA work
- Low level of cross-pod interaction
- Each pod has its own “process personality”, which presents challenges for clinic-wide implementation
Figure 3-1: Fraction of observed time spent managing prescriptions. Source – MIT direct observations.
• Many small IT problems generate rework and confusion

• Implementation of quality improvement initiatives, such as team huddles, is very inconsistent across pods and providers

• A large portion of non-visit patient contact takes place through the phone

• Low consistency between pods of assignment of tasks to roles

In addition to supporting prescription management improvement, the time study data provided insight into several potential avenues of future work (see Section 5.5 for details).

3.3 Methods

Quantifying the current state is necessary to determine how and where IMA is expending its resources, support development of alternative care delivery models, and provide baseline data for decision support tools to quantify the benefits/risks of any proposed changes.

The basic work unit for patient care within IMA is an appointment. Providers support appointments in half-day sessions while other staff support providers/patients/appointments in pairs of sessions (two sessions equals one work day). The observation data collection plan has two main targets: 1) follow patients at each step of their appointments, and 2) follow staff members through a complete session. This approach aims to capture the process flow from both a patient-centric and IMA operations standpoint. Sessions were used as the basic observation unit to match up with the scheduling system. Samples of the observation forms developed to capture data are provided in Appendix A and Appendix B. The major components of each observation are:

• Info Blocks: Identify the date/time/location/role/etc. of the observation

• Time: recorded to the nearest minute

• Location: physical location within the practice

• Staff present: What other practice members are interacting with the observed person
• Activity: What task the observed person was working on

• Notes: A free-text description of the task and any relevant information

The "Activity" fields were developed in consultation with IMA to provide a meaningful way to categorize the disparate activities conducted in the practice. Activities are first divided as "Visit" (V) or "Non-Visit" (NV). Visit activities correspond to work required to support a scheduled appointment, are associated with a patient physically present or planned to be present, and are associated with practice revenue under a fee-for-service payment system. Non-visit activities are everything else, including Population Management (PM), patient care not associated with an appointment (Intravisit care, or "Intra"), and administration. Detailed guidelines for assignment activities are printed on the back of the observation forms (included in Appendix A & B) and repeated here for convenience:

• Visit: Work associated with a specific billable/scheduled visit (past, present, or future)
  
  – Pre: Work to prepare for the patients visit
  
  – Treat: Working with a patient
  
  – Post: Work to follow-up on patients visits that already occurred
  
  – Oth: Anything else related to a particular visit (see the notes field for details)

• Non-visit: Work not associated with a specific billable/scheduled visit
  
  – PM: Population management (as defined by IMA role documents)
  
  – Intra: Care for patients not directly tied to a visit (letters, emails, calls, consultations, prescriptions, etc.)
  
  – Adm: Administrative work not related to any patient or population need
  
  – Oth: Anything else

Observers did not interact with patients and were not present in the exam room per current MGH Internal Review Board (IRB) guidelines. Individual staff members were followed at each step through a session. When appropriate, the time that patients entered or left
the exam room or interacted with observed roles was recorded. Role/pod combinations were chosen with guidance from IMA to provide a representative spread across the practice.

Completed forms were scanned and form data was transcribed for analysis. This dataset includes all of the information captured on the observation forms as well as the following derived data:

- Duration: The time, in minutes, for each line
- Rx-Related: A TRUE/FALSE flag for whether a line is related to prescription handling, determined by the text “rx” or “presc” in the notes field

### 3.4 Data Collection & Analysis

We initially proposed tracking two patients per session in ten different pods (twenty patients, ten sessions) with patients selected in concert with the MDs approximately 2–3 days ahead of the session. Identified patients would have been informed by the PSC at check-in using an IMA-approved script. After initial role observations it was clear that this approach was unnecessary, as most patient interactions were observed in the process of the role observations. As a result, only eleven dedicated patient observations were completed by recording the time that particular patients arrived in the practice and moved from one practice staff member to another. This data is identical to that captured while observing roles, and the 11 dedicated patient observations were combined with the role observations to cover a total of 77 appointments.

We initially proposed tracking one of each of the eight major patient-facing roles through a session in at least two different pods (three observations per role, twenty-four observations total, twenty-four sessions). As the observation effort progressed, the following changes were made with IMA concurrence:

- Patient Service Coordinator (PSC): Increased from 3 to 9 due to request from the Operations Manager and recognition that observing the PSC was the most effective way to see overall pod workflow
- Medical Assistant (MA): Increased from 3 to 4 due to scheduling
Table 3.1: Total role observations

<table>
<thead>
<tr>
<th>Role</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preceptor</td>
<td>3</td>
</tr>
<tr>
<td>Attending</td>
<td>9</td>
</tr>
<tr>
<td>Resident</td>
<td>2</td>
</tr>
<tr>
<td>Nurse Practitioner</td>
<td>3</td>
</tr>
<tr>
<td>Nurse</td>
<td>6*</td>
</tr>
<tr>
<td>Medical Assistant</td>
<td>4</td>
</tr>
<tr>
<td>Patient Service Coordinator</td>
<td>9</td>
</tr>
<tr>
<td>Total Observations</td>
<td>36</td>
</tr>
<tr>
<td>Total Duration (hh:mm)</td>
<td>127:12</td>
</tr>
</tbody>
</table>

* - In one case, two observers shadowed the same nurse at the same time. Thus, in the quantitative analysis \( N = 5 \) for nurse observations.

- Nurse: Increased from 3 to 6 to capture more data for Non-visit care activities
- Resident: Reduced from 3 to 2 due to similarity to attending observations and time constraints
- Attending: Increased from 3 to 9 to increase practice confidence in the validity of the attending data

The final distribution of role observations are provided in Table 3.1.

A total of three different observers were used to compile the observations. Initial observations were done in pairs to improve consistently between observers.

### 3.5 Time Study Results

This section discusses both the qualitative and quantitative findings from the time study. Observations highlighted several system design challenges that could be addressed to improve practice operations.

#### 3.5.1 Schedule Performance

As part of the provider observations, many patients interacted with the providers. Comparison of the times the providers entered and left the exam room with the schedule allows
Figure 3-2: IMA schedule performance. Delay – difference between scheduled and actual start time. Overrun – difference between scheduled and actual duration. Each marker corresponds to a single appointment, with like symbols corresponding to the same provider.
calculation of the appointment delay and overrun. The aggregate results are shown in Figure 3-2. Significant items include:

- A "perfect" scheduling system would cluster around the origin (0 delay, 0 overrun)
- The average appointment is both delayed and overrun, to high statistical significance
- There is clustering by provider: providers tend to be consistently on time or consistently late
- There is a statistically significant (P-value <= 0.05) positive correlation between overrun and delay. Once a provider is delayed, they are more likely to overrun and become further delayed for subsequent appointments.

3.5.2 Distribution Of Time Across Tasks

It is helpful to look at the observation data in terms of what fraction of each roles’ time is taken up by different tasks. See Section 3.3 for detailed task definitions. Figures (3-3) – (3-9) show the fraction of observed time spend on each task with one figure per role in standard boxplot format. Note that each fraction is the total of many very small (1-2 minute) tasks and there is wide variability within roles and across roles, consistent with the observation of significantly varying workloads and processes across pods and throughout the day.

**Figure 3-3:** Based on responsibilities we expect PSCs to have a high fraction of pre- and post-visit work and non-visit administration, and this is the case. However, the non-visit intra activity (patient care not associated with an appointment) is equally significant. This is driven primarily by PSC handling of clinical phone calls and faxes (prescriptions, referrals, etc.).

**Figure 3-4:** As expected based on responsibility, the medical assistants spend a high fraction of their time supporting visit activities. However, as with the PSCs, note the significant fraction of non-visit care (NV-Intra) time.

**Figure 3-5:** As expected, the nurses spend the majority of their time on non-visit care (primarily triage phone calls, follow-up phone calls, and prescription management).
However, note the extreme variability across the observed nurses, suggesting substantially different nurse workflows between pods.

**Figure 3-6:** Comparable to attendings and residents, nurse practitioners spend the majority of their time on patient treatment. However, as expected for their flexible role, the nurse practitioners spend a greater fraction of their time on intravisit care (NV-Intra).

**Figure 3-7:** As expected, residents spend the majority of their observed time actually treating patients. This pattern is very similar to attendings, as we would hope, but with smaller spread (likely due to the small sample size).

**Figure 3-8:** Similar to residents, attendings spend the majority of their time on patient treatment during a visit. However, note the higher pre- and post-visit load and larger spread compared to resident data.

**Figure 3-9:** Consistent with the observed preceptor workflow, preceptors spend the majority of their time on visits and post-visit consultation with residents.

### 3.6 Discussion

Based on the combination of direct observation, data logs for appointments and prescriptions, and derived metrics discussed in detail in Section 4.3, we identify prescription management as the first area to make a concrete intervention in the practice. Chapter 4 details the analysis, modeling, and proposed operational changes to address prescription management at IMA. In the course of performing the observations, the MIT team also noted several more general operational items of interest, which are provided for reference in Appendix D.
Figure 3-3: Time distribution by task – Patient Service Coordinator (PSC)

Figure 3-4: Time distribution by task – Medical Assistant (MA)
Figure 3-5: Time distribution by task – Nurse (RN)

Figure 3-6: Time distribution by task – Nurse Practitioner (NP)
Figure 3-7: Time distribution by task – Resident

Figure 3-8: Time distribution by task – Attending (MD)
Figure 3-9: Time distribution by task – Preceptor
THIS PAGE INTENTIONALLY LEFT BLANK
Chapter 4

Designing An Improved Prescription Management Model

4.1 Current Process & Resource Use

Based on the data in Chapter 3, prescription management represents a significant consumer of IMA resources and shows several opportunities for improvement, including: improvement of the yield of prescription request, reducing excess prescription requests, improving the number of prescriptions dealt with during appointments, and improving the overall workflow so that proper prescription requests are fulfilled in the most efficient way.

In order to quantify resource utilization in terms of Full-Time Equivalent (FTE) staff, we document the current state workflow, develop new operations metrics, and develop a network flow model of to quantify resource utilization in the current and recommended future states. The model shows the practice utilizes an average of 4.1 FTE to service non-appointment prescription requests. This estimate is consistent with the utilization noted during direct observation.

Direct observations and interviews with key practice members were used to generate the current state workflow (Figure 4-1). This simplified model neglects some niche request processes, such as direct emails to doctors or walk-ups, because these are workflow equivalent to modeled workflows (a direct email to a doctor is routed similarly to a web request, a walk-up is routed similarly to a voicemail request). It is important to note that, since each pod
Figure 4-1: IMA current state prescription workflow - single pod

Figure 4-2: IMA current state prescription workflow - all pods
is operationally independent, the actual practice workflow is highly parallel as illustrated in Figure 4-2. This allows us to model an "average pod" as one fourteenth of the total practice.

4.2 Operational Challenges

IMA staff are frequently interrupted and spend significant time refilling prescriptions. As expected from discussions and workflow, the prescription load falls heaviest on the nurses and medical assistants. Note that, due to the IRB restriction to not be in the room with patients, the NP/attending/resident fractions do not include prescription management that may have happened during an appointment. This did not influence the final recommendations because the process redesign is focused on the non-appointment prescription process.

Figure 3-1 shows the cumulative amount of time spent on prescriptions as a fraction of total observed time. This work typically occurred in many very short (1-2 minute) increments throughout a session. In addition to the time spend on handling the prescription, there is an associated cost of switching tasks so frequently (see Section 2.3.2 for relevant literature).

Figure 4-3 shows the aggregate observed prescription workflow. Note that no single pod was observed to execute all of the possible workflows; the complex nature of the workflow in figure 4-3 is a result of high variance in the prescription workflow between pods. Within any particular pod, the actual workflow also varied depending on the provider, drug, request path, and output path.

4.3 Metrics

In addition to the extensive time studies discussed Chapter 3, a full characterization of IMA resource utilization must include an analysis of the different input, processing, and distribution channels for prescriptions. Metrics to quantify operational performance of each channel are also developed.
Figure 4-3: Current state non-appointment prescription workflow - all pods
4.3.1 Prescription Volume

The prescription logs from the electronic medical record include the data for every prescription (patient, drug, dose, frequency, doctor, etc.), as well as a variety of useful metadata (person who actually created the prescription, distribution channel, pharmacy, etc.). Analysis of this data with respect to the known location of each doctor and nurse within the practice identifies, for each prescription, the corresponding pod and distribution channel. We further make the assumption, supported by direct observation and interviews with support staff, that prescriptions issued for a patient on the day of an arrived appointment were written by the doctor during face-time with the doctor. This allows us to separate prescriptions into appointment-based and non-appointment-based. The annual traffic for the entire practice by channel is shown in Figure 4-4. This can also be broken down by pod and prescription issuer as shown in Figure 4-5.

Prescription volume can also be assessed in terms of delivery channel, as shown in Figure 4-6. Electronic prescriptions, by necessity, are coded with the destination pharmacy and this data is captured in the prescription log. One retail pharmacy network is, by a wide margin, the largest filler of electronic IMA prescriptions. The second highest volume channel is hard-copy prescriptions handed or mailed to the patient; the final destination of these hard-
Figure 4-5: 2012 Total prescription volume, by pod, by role. 2010 and 2011 are qualitatively similar.
copy prescriptions is invisible to the practice, which increases the challenging of tracking prescriptions through fulfillment and ultimate drug adherence. If we assume that hard copy prescriptions are allocated among pharmacies in the same proportion as electronic prescriptions, the primary retail network is handling more than 40% of the total practice prescription volume. This is a reasonable assumption because the decision of how to deliver a prescription and which pharmacy to deliver it to are essentially independent; the OnCall system will support any delivery method for any pharmacy and patients can freely choose both the delivery method (excluding controlled substances) and the pharmacy. Due to this high volume, the pilot pharmacy intervention discussed in more detail in Section 4.6.1 was conducted with a major retail pharmacy network.
Table 4.1: Requested vs. issued prescriptions - normalized to 2012 total request volume

<table>
<thead>
<tr>
<th>Channel</th>
<th>Requested</th>
<th>Issued</th>
<th>Delta</th>
<th>Yield</th>
<th>Main Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web</td>
<td>3%</td>
<td>2%</td>
<td>&lt;1%</td>
<td>93%</td>
<td>Patients</td>
</tr>
<tr>
<td>Voicemail</td>
<td>8%</td>
<td>5%</td>
<td>3%</td>
<td>70%</td>
<td>Patients</td>
</tr>
<tr>
<td>Fax</td>
<td>47%</td>
<td>35%</td>
<td>12%</td>
<td>74%</td>
<td>Pharmacies</td>
</tr>
<tr>
<td>Appointments</td>
<td>42%</td>
<td>42%</td>
<td>-</td>
<td>100%</td>
<td>Patients</td>
</tr>
</tbody>
</table>

4.3.2 Non-Actionable Prescription Requests (Yield)

The prescription logs are only recorded when a prescription is actually entered into the medical system. There is currently no tracking of requests for prescriptions as an independent quantity even though this is the primary driver of workload to the Patient Service Coordinators and Medical Assistants. We therefore augmented the raw volume measurements discussed in Section 4.3.1 with direct observations of the inbound fax, voicemail, and web processes to determine how many prescription requests actually occur to generate the recorded issued prescriptions, the yield of the process. The results are shown in Table 4.1.

Yield: The ratio of issued to requested prescriptions for a particular request path

Of particular note is the high fraction of total volume moving through the fax process, the very high yield of the web process, and the low yield of the fax and voicemail processes. This suggests several qualitative drivers for operational improvement:

- Any improvement to the fax process has disproportionate impact due to its dominant volume

- The web process is the most effective but also the least utilized; IMA had already recognized this prior to this project and is currently implementing a plan to significantly increase the number of patients using the web portal

- The voicemail system is both low volume and low yield; there is opportunity to either increase the yield through system modification or reduce the volume by shunting patients to other, more efficient, input channels
4.3.3 Spillover

During direct observation, many support staff commented on a perceived high frequency of prescription requests happening immediately before and after appointments. This was somewhat unexpected, as the in-appointment prescription renewal process was perceived to be very efficient and doctors stated they made a concerted effort to make sure patients' prescriptions were up-to-date during any appointment. In order to quantify this observation, we developed a new metric, "Spillover":

Spillover: the percentage of prescriptions that occur within two days of a scheduled appointment (before and after), but not on the day of the appointment

As discussed in Section 4.3.1, prescriptions on the day of an arrived appointment are considered to have been issued during the appointment. The method to compute Spillover is as follows:

1. For every entry in the prescription log, identify the closest scheduled appointment in time for that patient

2. Calculate the time differential as \( \text{PrescriptionDate} - \text{AppointmentDate} \) (a negative value means the prescription was issued before the appointment)

3. Count all prescriptions with a time differential of \(-2, -1, 1, 2\) (not 0) days

4. Divide the count from Stem 3 by the total number of non-appointment prescriptions. The resulting value is "Spillover".

The spillover concept is illustrated graphically in Figure 4-7. Spillover was 8.5\% in 2012, with comparable values in 2010 and 2011. The impact of spillover is shown in Figure 4-8: spillover resulted in an extra 7,782 prescription requests going through the non-appointment prescription renewal process in 2012, or 4\% of the total prescription volume. The anecdotal reports obtained during observations are supported by quantitative analysis; there is a high volume of prescription activity surrounding each appointment. As a measure of effectiveness of the in-appointment prescription process, spillover provides a useful measure of how much unnecessary work the non-appointment renewal process performs to make up for prescriptions
that, for whatever reason, were not issued during the appointment itself. Additional research is required to determine the causes and addressability of any failures of the appointment-based prescription process.

4.3.4 Excess Chronic & Stable

Prescription renewals benefit, at least in theory, from a relatively high degree of predictability. The practice knows, via the electronic medical records, the prescribed drug therapy and active prescription duration for every patient, therefore it should be relatively straightforward to predict the incoming volume of renewal requests for at least certain classes of drugs and patients. Absent other constraints (controlled substances, dose titration, one-time therapy), the practice will always write a one year prescription. Although the practice does not currently monitor the list of outstanding prescriptions to predict when renewals should arrive, we hypothesize that patients on chronic medications at constant dosage should be renewing every 11–12 months (1.0–1.1 prescriptions per year). To test this hypothesis, we develop a
metric called “Excess Chronic & Stable”:

Excess Chronic & Stable: Average prescriptions per year per patient for each constant dose chronic medication prescribed for more than one year to that patient

The method to compute Excess Chronic & Stable is as follows:

1. Identify those classes of drugs typically prescribed chronically (multi-year therapies) at constant dosage. For this study, these were:
   - Anti-hypertensives
   - Anti-lipemics
   - Oral anti-diabetics
   - Thyroid simulators

2. Filter the prescription list for the drug families identified in step 1. As these are annual prescriptions, this list must span multiple years for the calculation to be valid. 2010-2012 were used for this study.
3. Filter the resulting list for patients who have multiple prescriptions. Patients with only one are assumed to have either started or stopped therapy within the window of interest, hence do not have a stable prescription pattern.

4. Filter the resulting list for prescription renewals at the same dose. Dose changes may indicate titration, revised lab results, or other clinical actions that prevent the prescription pattern from being stable.

5. Calculate the average number of prescriptions per patient per year for the chronic & stable prescriptions identified in step 4. This value defines “Excess Chronic & Stable.”

If our hypothesis is correct, “Excess Chronic & Stable” should have a value of 1.0 – 1.1, signifying patients renewing as their prescriptions run out at month 11 or 12. This expected pattern is illustrated in the upper half of Figure 4-9. However, processing of actual logs shows a value of 1.8, equivalent to renewing every 7 months and shown in the lower half of Figure 4-9 (t-test P-value < 0.0001 against expected value of 1.1). This implies that a significant fraction of prescription requests are not clinically necessary; the patient has an active prescription with available refills but, for whatever reason, the patient is requesting a new prescription well ahead of the expiration of the old one. This generates additional workload on the practice for no clinical benefit and increases the likelihood of a prescription linking error at the pharmacy. The volume generated by these excess requests is shown in Figure 4-10. In 2012, prescriptions for chronic & stable drug therapies represented 20% of the total volume, while almost half of that (8% of total volume) were excess. Reduction of excess renewal requests represents a significant opportunity to reduce total prescription volume without any impact to patient care. In addition, as this definition of chronic & stable is conservative, it is likely that a significant fraction of the remaining 80% of total prescriptions are also renewed early. Thus, we expect 8% to be a lower bound for the amount of excess renewals flowing through the system.
Figure 4-9: Metric schematic – excess chronic & stable prescriptions

Figure 4-10: 2012 Total excess chronic & stable prescriptions
4.4 Rationale For Proposed Solution

A network flow model is built utilizing the observed workflow for the current state. Adding nodes and arcs to the model allows alternative workflows to the overlaid on the current state workflow, estimating volume and resource requirements for each potential prescription workflow design.

The final network flow model is shown in Figure 4-11. The black network presents the current state while the gray network represents the proposed future state. Separate processing and wait times are assigned to each arc. The model is formulated as a linear program that attempts to minimize total processing time subject to flow constraints and traffic splits that correspond to the observed metrics discussed in Section 4.3. In both the current state and future state models, the constraints are binding, such that the feasible region is a single point. As such, the model could also be developed as a series of linear equations and solved directly. However, the network flow/optimization approach provides the option to introduce arcs (workflows) that may expand the feasible region, preserving flexibility to find optimal workflows during the iterative phase of experimenting with alternative process designs.

4.5 Modeling

The model takes, as input, the following:

**Patients:** Total patients, fraction of patients on medications, average prescriptions per patient per year, and daily standard deviation of prescription demand

**Request Channels:** The fraction of requests coming by the four primary channels (Appointment, Web, Voicemail, & Fax)

**Metrics:** The yield of each request channel (Section 4.3.2), spillover prescriptions (Section 4.3.3), and chronic & stable request ratios (Section 4.3.4)

**Times:** Task times based on the direction observation results (Section 3.4) and wait times estimated with a dedicated discrete event simulation model
Figure 4-11: Network flow model schematic
The model also includes parameters to account for the impact of various potential interventions and redesigns:

**Interventions:** Staffing level for the central prescription group, how many pods worth of traffic to run through the centralized group, and the expected specialization and process time savings in the central group.

Based on these inputs, the model calculates the input traffic of prescription requests by channel (source nodes), and the output traffic of issued prescriptions (sink nodes). Metrics, such as yield and spillover, define the ratio of traffic on different arcs. These calculations are expressed as linear constraints and combined with the flow constraint that net flow on non-source/sink nodes be zero. The cost matrices use the task and wait times to assign a cost to each arc. The optimization model is built in Excel for easy portability and solved using OpenSolver (the number of decision variables exceeds the capability of the native Excel solver). Solving the model to minimize total processing time provides the traffic on each arc. Total processing time on all arcs measures the FTE expended to serve the prescription volume, while total weighted average wait+process time along all source-to-sink paths provides the average expected cycle time.

The model also assumes that each pod’s traffic can be modeled as one of fourteen independent identically distributed random variables that sum up to the observed practice volume and standard deviation. As a result, the model can predict the mean and standard deviation of volume for the separate pods and for the centralized prescription group.

The model outputs:

**Volume:** Total prescriptions issued, total non-actionable requests, total spillover and total excess prescriptions issued

**Variability:** Coefficient of variation for the practice, each pod, and the central group, including average, +1σ, and +2σ daily demand and FTE required

**Savings:** Average cycle time and wait time, including improvement against the current state baseline

**Resources:** Total FTE, FTE by role, and utilization by role
4.5.1 Model Formulation

$x_{ij}$: Number of prescriptions flowing from node $i$ to node $j$

$N_{ij}$: Node adjacency matrix, $N_{ij} = 1$ if there is an arc from node $i$ to $j$, 0 otherwise

$C_{ij}^w$: Wait time, per prescription, from node $i$ to node $j$

$C_{ij}^p$: Processing time, per prescription, from node $i$ to node $j$

$M$: "Infinity", larger than the largest value on any arc, $= 10^6$

$F$: Indicator for use of central clinical processing, $F = 1$ when central clinical processing is used, 0 otherwise

$\mathcal{A}$: The set of source (supply) nodes

$\mathcal{B}$: The set of sink (demand) nodes for complete prescriptions

$\mathcal{N}$: The set of arcs from central administrative scrub node(s) to central clinical processing node(s)

$\mathcal{E}$: The set of arcs that carry excess prescriptions

$\mathcal{W}$: The set of sink nodes for waste prescriptions

$\mathcal{Y}$: The set of arcs that carry spillage prescriptions

Decision Variables: $x_{ij}$

Minimize: $\sum_{ij} C_{ij}^p x_{ij}$ subject to:

$$\sum_j (N_{ij} x_{ij}) = \text{Requests}_i \quad \forall i \in \mathcal{A}$$

$$\sum_i (N_{ij} x_{ij}) = \text{Outputs}_j \quad \forall j \in \mathcal{B}$$

$$\sum_i (N_{ij} x_{ij}) = \text{Waste}_j \quad \forall j \in \mathcal{W}$$
\[
\sum_{ij} (N_{ij}x_{ij}) = \text{Excess}_{ij} \quad \forall ij \in \mathcal{E}
\]
\[
\sum_{ij} (N_{ij}x_{ij}) = \text{Spillage}_{ij} \quad \forall ij \in \mathcal{M}
\]
\[
x_{ij} \leq (1-F)M \quad \forall ij \in \mathcal{N}
\]
\[
\sum_{i} (N_{ij}x_{ij}) - \sum_{i} (N_{ji}x_{ji}) = 0 \quad \forall j \in \mathcal{A}, \mathcal{B}, \mathcal{W}
\]

As formulated, the model has \(15^2 = 225\) decision variables, one for each possible arc, and 27 constraints. In the final design, due to the relatively sparse network, there are only 29 active decision variables.

None of the arcs or nodes in this formulation are capacity limited as no portion of the current IMA workflow saturates purely due to prescription load. A network including capacity limits might be appropriate for future models that attempt to incorporate other workflows that compete for practice resources, or if the prescription utilization of any node in the workflow approaches 100%.

The model was validated by running it against actual input values from 2012 with no process improvements (no centralization or specialization benefits) and comparing the model outputs of predicted volumes to the actual logs of prescription outputs in 2012. The predicted FTE per pod was also compared to the fraction of time spent on prescription management found during direct observation (Figure 3-1), with essentially equivalent results after accounting for pod-to-pod variation in workflow and load.

Screen shots of the as-implemented model inputs, adjacency matrix, and outputs are provided for reference in Appendix E.

### 4.5.2 Modeling Results

Modeling shows that the practice currently devotes approximately 4.1 FTE to non-appointment prescription management. This is notable for several reasons:

- This is the first time the practice has a quantitative estimate for the resource consumption of any dispersed parallel task
- IMA is large enough that the prescription management task significantly exceeds one FTE, therefore it is at least possible to make this a dedicated position. Most primary care practices are too small to have dedicated staff for prescriptions but IMA has an opportunity to leverage their scale to utilize specialized and dedicated roles.

- Based on bench-marking against other MGH primary care practices, IMA’s prescription management process is resource intensive. The next largest practice, Bullfinch Medical Group, performs prescription management for 20,000 patients (2/3 of IMA’s patient population) using 1.0 – 1.5 FTE.

Modeling also shows that pooling all pods into a central prescription management group would reduce daily coefficient of variation from 0.655 (per pod) to 0.175, assuming 14 average and equal pods working in parallel (Figure 4-12). This compares favorably to the coefficient of variation computed from the historic prescription logs, with an average pod coefficient of variation of 0.49 dropping to 0.19 with full pooling for 2012 data. In FTE terms, this implies that a +2σ demand day (covering 98% of all days) would require an extra 0.7 FTE in an individual pod, a 130% increase, but only an extra 1.4 FTE for the entire practice, a 35% increase. The revised workflow with centralization is illustrated in Figure 4-13.

Centralization, by itself, does not imply a reduction in total workload since all the same work must still be done. As a result, the model estimate of 4.1 FTE for the current state also applies to the central prescription management group absent any other improvement. However, the variability reduction of pooling means that a centralized prescription group does not require as many resources to handle higher-than-average workload days. It is also well established in operations literature that segmenting a particular process away from others allows employees to focus and specialize, reducing task times and improving throughput, simplifying change implementation, and easing troubleshooting. We undertook a series of pilot studies to quantify how much improvement could be expected at IMA.
Figure 4-12: 2012 daily prescription volume distribution and coefficient of variation, by pod
4.6 Implemented Interventions

Pilot interventions were conducted around reducing fax volume, improving voicemail yield, and redesigned workflows to estimate how much the model inputs could be improved.

Incorporating the pilot results into a centralized prescription management process shows that $2.0 - 2.5$ FTE should be able to handle the entire non-appointment prescription load for the practice, a 37% decrease in resource utilization. The central prescription management process also provides the foundation for several future interventions necessary to implement proactive medicine tracking and management for patients.

4.6.1 Prescription Linking At The Pharmacy

As discussed in Section 4.3.2, a significant fraction of total prescription requests are due to non-actionable faxes. These typically take the form of duplicate renewal requests from pharmacies: they request renewal of a prescription that has not yet expired or that was recently renewed. MGH Back Bay, a smaller primary care practice, had previously worked
with a major retail pharmacy network to successfully reduce the number of duplicate faxes but was unsure what exact fixes the pharmacy had applied or whether it would be applicable to a practice the size of IMA.

As the provider of 30-40% of IMA’s total prescription fulfillment, any improvement at the primary retail pharmacy network could make a meaningful impact to overall practice prescription volume. We contacted one major retail network to examine the issue further. They have a centrally administered IT system but operational execution of prescriptions, including renewals, is performed at the store level. They were very interested in determining the source of duplicate faxes as this represents additional non-value-added work for their stores.

Retail stores in this network will initiate a renewal request fax if a patient contacts them for a refill and their current therapy has no remaining refills. In order for patients to have continuous therapy, the pharmacy links new prescriptions to old ones. This linking processes is performed manually at the store level. If a store fails to link the prescription, the new prescription goes into a “Hold” status. It is available to be activated when the patient requests a refill but the pharmacy technician must locate and activate the held prescription; it is often simpler for the pharmacy technician to request a renewal than to search for a potentially held prescription. The central office hypothesized that the stores were not linking prescriptions and simply relying on renewal faxes.

In order to test this theory, the pharmacy network offered to take a list of IMA doctors, identify all patients under those doctors using their pharmacies, then intervene with the affected pharmacies to manually link all held prescriptions in their system. Since the central office has no visibility into individual store actions with respect to fax requests, this had to be a cooperative effort between the central office and IMA to provide sufficient visibility to both sides of the process.

We provided a list of all attending doctors and related identifying information for IMA Pods 2 & 3 to the central office and began monitoring the fraction of faxes from that pharmacy network that were labeled as duplicates, defined as no IMA action required because an appropriate prescription had already been issued. During the control period ahead of the linking implementation, the mean fraction of duplicate faxes was 23%. After the intervention
the mean duplicate fraction was 15%, a 35% reduction in duplicate fax volume suggesting a significant shift in duplicate faxes (P-value < 0.05).

This retail pharmacy network plans to implement automated prescription linking across their stores starting in mid-2014. We expect a comparable reduction in duplicate faxes from this network for all IMA pods as this implementation rolls out, reducing total fax volume to the practice by > 9000 faxes per year.

This pilot also shows that meaningful improvement in prescription request volume is possible through better pharmacy coordination; under the prior IMA prescription workflow this was effectively impossible, as such coordination would involve approximately 28 people, none of whom had primary responsibility or accountability for the efficiency of the overall process. By segmenting and centralizing the prescription group, the process of pharmacy coordination is much simpler and we can reasonably expect comparable improvements from all of the major pharmacy partners.

4.6.2 Voicemail Renewal Line Retirement

As discussed in Section 4.3.2, voicemails are a relatively low traffic and low yield request process. They also take significantly longer to process (mean 4m13s) compared to faxes or web requests (mean 2m27s) due to having to transcribe data from the voicemail to the renewal form at the speed of the patients' speech and constrained by the voicemail system script. Staff and patients both expressed dissatisfaction with the voicemail system due to a perceived difficult interface, time consuming input process, and frequent patient reports of unfilled prescriptions.

Investigation of the issue through interviews and direct observations revealed the following:

- Both pharmacies and the practice would prefer that patients directly contact the pharmacy, rather than the practice, when a renewal is required. As discussed in Section 4.3.4 and 4.6.1, the practice frequently receives renewal requests significantly before the prior prescription has expired. If the pharmacy linking procedure is functioning properly, the pharmacy can identify if a renewal is required and does not need to
contact the practice unless the current prescription has actually expired.

- Patients frequently do not have the information necessary to complete the prompts in the voicemail script, resulting in incomplete voicemail messages or hangups.

- The voicemail system itself occasionally loses messages. A parallel research effort by a Harvard Medical School student researched this issue by placing test messages at known times into the system and then tracking whether they were picked up by the relevant Medical Assistants. Approximately 10% of messages were lost without explanation, validating the patient claim of unfilled prescriptions. As of this writing, MGH telecom services group is still attempting to identify the root cause of the missing messages.

- The prescription renewal lines use the term “refill” and “renewal” interchangeably despite the fact that these clinically very different. A refill is a quantity of drugs to be dispensed at a time; one prescription can authorize multiple refills. A renewal is a new prescription. The practice hypothesized that this ambiguous terminology may be leading patients to call the renewal line to leave a voicemail when they only need a refill, which should require no action from the practice. Pharmacies identified that they are aware of this ambiguity within the general patient population and this is one of the reasons they prefer that patients contact the pharmacy first, rather than going directly to the practice.

After extensive consultation with nurses, medical assistants, and the IMA operations manager, a new voicemail script was developed to replace the existing script, making the following changes:

- All references to “refill” are consistently changed to “renewal”

- Patients are encouraged to contact their pharmacy first, or use the web interface, and only leave a voicemail if their request cannot be serviced by any other channel

- Patients are prompted with the information that will be required and given the opportunity to pause and collect the required data prior to recording their request
• The information prompts are condensed to require less button pushing by the patient and speed up transcribing for the medical assistant

Once implemented, these changes are expected to reduce the volume of voicemail requests (by shunting that traffic to the more efficient fax and web systems) and improve the yield of the remaining voicemails (by improving the information contained in each message). As of writing, the script is currently being implemented by MGH telecom services. A schematic of the revised script is provided for reference in Appendix C.

The revised voicemail system will serve as a bridge; the ultimate goal is to retire the voicemail system completely and send all calls directly to the Patient Service Coordinators, where all other calls are currently handled today. The PSCs are in the best position to interactively capture the required information and route it to the appropriate staff member for action.

4.6.3 Prescription Renewal Workflow

As discussed in Section 4.5, centralizing prescription management provides benefits in terms of reduced variability through pooling; and improved specialization and cycle times through segmentation. However, centralization requires two significant changes from the current state workflow shown in Figures 4-1 and 4-2:

Signature compliance: The current state workflow relies on nurses to issue the prescription to the pharmacy. However, only doctors and nurse practitioners are licensed to issue prescriptions. IMA currently accommodates this requirement by having the doctor/NP cosign the medication note in the electronic medical record system but this cosign can occur days or weeks after the prescription is issued. MGH has identified that all primary care practices will need to transition to a process where a physician’s prescription order is provided before the prescription is sent to the pharmacy or patient.

Standard review: Under the current workflow, the degree of review performed by the Medical Assistant and Nurse varies significantly from pod to pod and from provider to provider. Each localized team has developed their own unique division of review tasks
with little coordination between pods. A central model must use a uniform process in order to capture the benefits of centralization, so a standard review that is acceptable to all providers must be included.

In order to accommodate these requirements the future state workflow (Figure 4-13) includes a standard review, followed by clinical message to the provider for prescription approval prior to releasing the prescription to the pharmacy. In order to prove that this workflow was feasible, two pilots were conducted in the following manner:

1. Begin holding prescriptions from two pods at noon the day prior to the pilot, in order to build an artificial backlog of prescriptions for processing

2. On the morning of the pilot, assemble a team of two medical assistants, two providers, and an on-call nurse, each with a computer

3. Process the backlogged prescriptions using the medical assistants serving the role of the central prescription group, the providers responding to the medical assistants via clinical messaging, and the nurse providing clinical support to the medical assistants as necessary

4. Record any discrepancies found during each prescription and update the draft standard work document as necessary

The first pilot focused on non-controlled substances with a focus on determining the appropriate level of detail required from the medical assistant review to enable the doctor to quickly assess the appropriateness of the renewal request and rapidly respond. This pilot demonstrated that clinical messaging could support the planned interaction and that the existing standard review document was adequate to provide sufficient information to the physician.

The second pilot focused on controlled substances and showed that the fundamental workflow developed during the first pilot would also work for controlled substances. This pilot also exposed several discrepancies in nuanced interpretation of the controlled substance rules between the pods and weaknesses in the existing controlled-substance checklist; changes were incorporated into the standard work and checklist to remedy these discrepancies.
The pilots demonstrate that the proposed workflow to support the central prescription management group is practical and allows prescription processing at a comparable rate to the current workflow. The pilots also serve as a powerful motivator for change within the practice by putting a practical face on research that had, until this point, been primarily theoretical and analytic.
Chapter 5

Recommendations & Conclusions

5.1 Results

Based on the data, analysis, and modeling detailed in Chapter 3 & 4, we demonstrate that:

- The current state workflow is highly parallel and does not take advantage of IMA’s scale.

- The prescription flow can be modeled as a network flow problem to estimate the total resources required to serve non-appointment prescriptions: 4.1 FTE. This result agrees well with direct observation of practice operations.

- Pooling the prescription workflows between pods should reduce coefficient of variability of daily demand from 0.49 to 0.19 based on 2012 prescription logs.

- Pooling the prescription flows into a central prescription management group and segmenting this workflow from other practice operations will allow staff to develop stable, standard, dedicated workflows and begin to address chronic prescription issues: spillover prescription requests, excess prescription requests, and non-actionable prescription requests.

- Coordination with pharmacies to properly link renewed prescriptions can reduce duplicate (non-actionable) fax renewal requests by > 35%.
• Voicemail is a sub-optimal method of requesting renewals from the point of view of the practice, pharmacies, and patients. Use of the voicemail system should be minimized or eliminated; until it can be eliminated, revisions to the voicemail script to reduce frequent patient errors can help increase voicemail yield.

• A centralized workflow where medical assistants perform a standardized review of each prescription request, followed by provider confirmation via clinical messaging, is practical and effective.

As the central prescription group becomes established and stable, we expect the following future state is attainable:

• Daily variability reduces > 50% due to pooling

• Fax yield increases to 87% due to improved prescription linking consistent with the CVS pilot

• Voicemail renewal lines are closed

• 50% of excess and spillover prescriptions are eliminated due to dedicated employees with a mandate to correct process errors

• Task time decreases 30% due to specialization enabled by segmentation

Under these assumptions, our model predicts that total resource use to serve non-appointment prescriptions will drop from the current 4.1 FTE to 2.0-2.5 FTE, a 37% reduction. At the same time, we expect improved patient care due to faster cycle times, less errors, and better drug adherence.

5.2 Performance Metrics

Currently, none of the data sets or computations used to generate the metrics discussed in Chapter 3 can be run continuously or in real-time due to IT system limitations. Prescription and appointment logs can be run on an ad hoc basis at short (< 5 day) lead time, followed by offline processing of the data to generate updated metrics. However, the PCOI patient
processing is run annually against a much larger third-party dataset, therefore any analysis based on patient characteristics such as comorbidities or practice loyalty can only be run after each calendar year.

We propose rerunning the available metrics (volume, variability, spillover, excess chronic & stable, and yield) after the first full month of operation of the central prescription group. Relative to the current state, we expect the changes shown in Table 5.1.

As the practice does not currently track any prescription-related patient outcomes, there are no defined patient outcome metrics. There is opportunity for future research to tie existing drug-related metrics, such as ED admissions or health-outcomes for drug-managed conditions, back to prescription management processes.

### 5.3 Pooling & Segmentation

As is generally proven in the operations literature, pooling will lead to reduced variability. Analysis of past data and modeling both show an expected 50% reduction in daily coefficient of variation by pooling the full practice prescription traffic into a central group.

Segmenting this group’s workflow away from other practice processes provides the opportunity to reduce interruption to patient-facing staff and allow dedicated prescription management staff to specialize and standardize to a degree that is impossible under the current workflow.

This process design also lays the foundation for any future change management with respect to prescriptions, as it reduces the amount of coordination for any change from 28
5.4 Proactive Medicine Management

Prescription renewal is inherently predictable to a large extent; the current system functions entirely by patient-pull primarily because the practice does not have the resources or processes to take advantage of the predictability inherent in the prescription data in the electronic medical record.

We hope that, once the process improvements identified by this research are implemented, the next step will be to experiment with proactive prescription management. As one simple example, the algorithms developed during this research could generate a list of patients expected to have a prescription renewal for particular drugs in any given time period (e.g. the coming week). A central prescription management group could then proactively issue those prescriptions and notify the patient, rather than waiting for patient contact and potentially missing a drug non-adherence.

Once the capabilities are in place to proactively identify what patients should receive which prescriptions when, and a dedicated staff is available to enact such policies, many exciting options become available for future improvements in prescription management.

5.5 Future Work

In addition to implementation and monitoring of the recommended prescription management system, the data analysis and observations inform several possible productive avenues of future research:

5.5.1 Scheduling

Rescheduling

A significant number of PSC activities involve scheduling and rescheduling appointments. Approximately 40% of booked appointments never arrive and the recovery rate (ability to
re-fill a canceled slot) is low. In theory, the schedule system provides tools to manage recovery rate via the waitlist function but observations showed that this function is rarely employed by the PSCs due to perceived lack of utility and user-unfriendliness of the IDX implementation. Interventions that can improve schedule utilization would lower the PSC workload and improve patient access by opening up provider schedules.

**Appointment Duration**

The average appointment is both delayed and overrun (see Section 3.5.1 for details). Observations showed poor correspondence between the schedule and actual appointment time, with predictable trends based on provider and patient. In virtually all cases, when asked, providers could correctly predict which patients would overrun. This suggests the possibility for a predictive model to better match appointment duration to patient needs.

### 5.5.2 Capacity Planning

**Adjustment For Patient Complexity**

There is a significant discrepancy between medical complexity (current computed by Risk Adjusted Panel Score) compared to “patient complexity” (how much effort and intensity is required by the provider to treat the patient). As a result, practice resources allocated based purely on medical complexity (e.g. case managers) may not correspond well to the actual practice resources required to care for the patient. A medically complex patient might be straightforward if they are also very proactive in their care, while a medically simple patient with significant psycho-social needs may take up significant practice resources for both their visits and intravisit care.

**Capacity Analysis - Panels**

IMA has been asked to increase panel sizes while most attendings have closed panels. There is currently no model to allocate new patients or to take account of the various metrics of patient complexity or current resource utilization. IMA will require a mechanism to allocate new patients that optimally balances available practice capacity.
Capacity Analysis - Support Staff

Observations show a consistent sentiment within pods that they are either under- or (rarely) over-staffed with respect to various support roles. Quantitative data supports this sentiment. For example, preliminary staffing ratio analysis shows that, although the overall provider to medical assistant ratio is approximately 1.0–1.75 (depending on the session), the range within any particular pod can be 0.0–3.0. This is supported by the role observations, where MAs sometimes had no providers on session and at other times had three. There appears to be a significant opportunity to alter the MA:provider mapping (either through provider schedule changes or MA reassignment) to level out the support ratio.
Appendix A

Role Observation Form
### IMA Time Study – Role Observation Form

<table>
<thead>
<tr>
<th>Start Time</th>
<th>Location</th>
<th>Staff Present</th>
<th>Activity</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WR ER Lab Desk</td>
<td>MD NP RN MA PSC Oth</td>
<td>Visit: Pre Treat Post Oth Non-Visit: PM intra Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IT Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR ER Lab Desk</td>
<td>MD NP RN MA PSC Oth</td>
<td>Visit: Pre Treat Post Oth Non-Visit: PM intra Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IT Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR ER Lab Desk</td>
<td>MD NP RN MA PSC Oth</td>
<td>Visit: Pre Treat Post Oth Non-Visit: PM intra Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IT Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR ER Lab Desk</td>
<td>MD NP RN MA PSC Oth</td>
<td>Visit: Pre Treat Post Oth Non-Visit: PM intra Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IT Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR ER Lab Desk</td>
<td>MD NP RN MA PSC Oth</td>
<td>Visit: Pre Treat Post Oth Non-Visit: PM intra Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IT Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR ER Lab Desk</td>
<td>MD NP RN MA PSC Oth</td>
<td>Visit: Pre Treat Post Oth Non-Visit: PM intra Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IT Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR ER Lab Desk</td>
<td>MD NP RN MA PSC Oth</td>
<td>Visit: Pre Treat Post Oth Non-Visit: PM intra Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IT Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR ER Lab Desk</td>
<td>MD NP RN MA PSC Oth</td>
<td>Visit: Pre Treat Post Oth Non-Visit: PM intra Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IT Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR ER Lab Desk</td>
<td>MD NP RN MA PSC Oth</td>
<td>Visit: Pre Treat Post Oth Non-Visit: PM intra Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IT Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR ER Lab Desk</td>
<td>MD NP RN MA PSC Oth</td>
<td>Visit: Pre Treat Post Oth Non-Visit: PM intra Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IT Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR ER Lab Desk</td>
<td>MD NP RN MA PSC Oth</td>
<td>Visit: Pre Treat Post Oth Non-Visit: PM intra Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IT Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR ER Lab Desk</td>
<td>MD NP RN MA PSC Oth</td>
<td>Visit: Pre Treat Post Oth Non-Visit: PM intra Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IT Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR ER Lab Desk</td>
<td>MD NP RN MA PSC Oth</td>
<td>Visit: Pre Treat Post Oth Non-Visit: PM intra Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IT Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR ER Lab Desk</td>
<td>MD NP RN MA PSC Oth</td>
<td>Visit: Pre Treat Post Oth Non-Visit: PM intra Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IT Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR ER Lab Desk</td>
<td>MD NP RN MA PSC Oth</td>
<td>Visit: Pre Treat Post Oth Non-Visit: PM intra Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IT Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR ER Lab Desk</td>
<td>MD NP RN MA PSC Oth</td>
<td>Visit: Pre Treat Post Oth Non-Visit: PM intra Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IT Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR ER Lab Desk</td>
<td>MD NP RN MA PSC Oth</td>
<td>Visit: Pre Treat Post Oth Non-Visit: PM intra Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IT Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR ER Lab Desk</td>
<td>MD NP RN MA PSC Oth</td>
<td>Visit: Pre Treat Post Oth Non-Visit: PM intra Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IT Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR ER Lab Desk</td>
<td>MD NP RN MA PSC Oth</td>
<td>Visit: Pre Treat Post Oth Non-Visit: PM intra Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IT Other</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Observer Name:**

---

Figure A-1: Role observation form, page 1

Property of MGH IMA – DO NOT RECORD PI/PHI ON THIS FORM – Apply PHI controls

Page of 84
Time Info:
Date: The date of the observation
Block: Whether this is a morning or afternoon block (circle one)
Seeing Pts?: Whether the role is scheduled to be seeing patients during this block (circle one)

Role Info:
Preceptor: An attending working on supervising residents (this does not include an attending seeing their own patients)
Attending: An attending working on their own patients
Pod: The hemipod where the person is working (i.e. if in “2” put “2”, not “2/3”)
Name: The person’s name (First Initial. Last Name)
FTE: The amount of Full-Time-Equivalent that the person typically works (can be obtained from Arlene if unknown)

Observations:
Start Time – The starting time (hh:mm:ss) for that line. No end time is needed (it’s the start time of the next line).
Location – Where, physically, the person is (circle one):
   WR: Waiting Room
   ER: Exam Room (record which in the notes)
   Lab: Blood or EKG lab (record which in the notes)
   Desk: The person’s normal workspace (office/MA desk/etc.)
   IT: In Transit (moving from one location to another)
   Other: Anything else (record where in the notes)
Staff Present – All practice members present with the person being observed (circle as many as needed).
   If multiple of one role, put a number by the role:
   MD: Attending
   R: Resident (if an MD is present and NOT the preceptor, put that in the notes)
   NP: Nurse Practitioner
   RN: Nurse
   MA: Medical Assistant
   PSC: Patient Service Coordinator
   Oth: Other (record who in the notes). Do not record people who came with the patient (family members, etc.)
Activity – What the person is doing:
Visit: Work associated with a specific billable/scheduled visit (past, present, or future)
   Pre: Work to prepare for the patients’ visit
   Treat: Working with a patient
   Post: Work to follow-up on patients’ visits that already occurred
   Oth: Anything else related to a particular visit (record in the notes)
Non-visit: Work not associated with a specific billable/scheduled visit
   PM: Population management (as defined by IMA role documents) – when in doubt, ask
   Intra: Care for patients not directly tied to a visit (letters, emails, calls, consultations, prescriptions, etc.)
   Adm: Administrative work not related to any patient or population need
   Oth: Anything else
Notes – Free text to elaborate on anything

Any time any of the observation fields change, you should start a new line. The time difference is then the duration for the prior line.

Figure A-2: Role observation form, page 2
Appendix B

Patient Observation Form
<table>
<thead>
<tr>
<th>Start Time</th>
<th>Location:</th>
<th>Staff Present:</th>
<th>Activity:</th>
<th>Notes:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WR Vitals ER</td>
<td>MO R NP RN MA PSC Oth</td>
<td>Wait Treat Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR Vitals ER</td>
<td>MO R NP RN MA PSC Oth</td>
<td>Wait Treat Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR Vitals ER</td>
<td>MO R NP RN MA PSC Oth</td>
<td>Wait Treat Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR Vitals ER</td>
<td>MO R NP RN MA PSC Oth</td>
<td>Wait Treat Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR Vitals ER</td>
<td>MO R NP RN MA PSC Oth</td>
<td>Wait Treat Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR Vitals ER</td>
<td>MO R NP RN MA PSC Oth</td>
<td>Wait Treat Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR Vitals ER</td>
<td>MO R NP RN MA PSC Oth</td>
<td>Wait Treat Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR Vitals ER</td>
<td>MO R NP RN MA PSC Oth</td>
<td>Wait Treat Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR Vitals ER</td>
<td>MO R NP RN MA PSC Oth</td>
<td>Wait Treat Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR Vitals ER</td>
<td>MO R NP RN MA PSC Oth</td>
<td>Wait Treat Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR Vitals ER</td>
<td>MO R NP RN MA PSC Oth</td>
<td>Wait Treat Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR Vitals ER</td>
<td>MO R NP RN MA PSC Oth</td>
<td>Wait Treat Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR Vitals ER</td>
<td>MO R NP RN MA PSC Oth</td>
<td>Wait Treat Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR Vitals ER</td>
<td>MO R NP RN MA PSC Oth</td>
<td>Wait Treat Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR Vitals ER</td>
<td>MO R NP RN MA PSC Oth</td>
<td>Wait Treat Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR Vitals ER</td>
<td>MO R NP RN MA PSC Oth</td>
<td>Wait Treat Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR Vitals ER</td>
<td>MO R NP RN MA PSC Oth</td>
<td>Wait Treat Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR Vitals ER</td>
<td>MO R NP RN MA PSC Oth</td>
<td>Wait Treat Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR Vitals ER</td>
<td>MO R NP RN MA PSC Oth</td>
<td>Wait Treat Adm Oth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WR Vitals ER</td>
<td>MO R NP RN MA PSC Oth</td>
<td>Wait Treat Adm Oth</td>
<td></td>
</tr>
</tbody>
</table>

Figure B-1: Patient observation form, page 1
Pt Info:
Patcom #: Unique identifier for each appointment in IDX. Available from the PSC and on the schedule report. Also recoverable after observation using provider name/time.
Notified of wait time?: Was the patient notified by a member of the practice (typically the MA) of the expected wait time? This is follow-up tracking for a prior IMA time study.

Time Info:
Date: The date the observation is made
Appt. Time: The scheduled time of the appointment
Sch. Duration: The scheduled duration of the appointment

Appt. Info:
Tick off the type of the appointment that was scheduled (available from the PSC). If the content of the appointment changes, that will be captured in the observation notes.

Provider Info:
Pod: The number of the hemipod (e.g. if you’re in 2 put “2”, not “2/3”)
Name: The provider’s name (First Initial. Last Name)
Type (circle one) – The type of provider that was scheduled:
  MD: Attending
  R: Resident
  NP: Nurse Practitioner
  RN: Nurse

Observations:
Start Time – The starting time (hh:mm:ss) for that line. No end time is needed (it’s the start time of the next line).
Location – Where, physically, the patient is (circle one):
  WR: Waiting Room (includes the PSC desk)
  Vitals: The vital signs measurement area (in the back hallway for most pods)
  ER: Exam Room (record which in the notes)
  IT: In Transit (moving from one location to another)
  Lab: Blood or EKG lab (record which in the notes)
  Oth: Anything else (record where in the notes)
Staff Present – All practice members present with the patient (circle as many as needed). If multiple of one role, write the number beside the circle and capture that in the notes:
  MD: Attending
  R: Resident (if an MD is present and NOT the preceptor, put that in the notes)
  NP: Nurse Practitioner
  RN: Nurse
  MA: Medical Assistant
  PSC: Patient Service Coordinator
  Oth: Other (record who in the notes). Do not record people who came with the patient (family members, etc.)
Activity – What the patient is doing:
  Wait: Waiting (i.e. not interacting with any member of the practice)
  Treat: Receiving treatment (physical or verbal) from any member of the practice
  Adm: Administrative work (appointment scheduling, pre- or post-visit paperwork, insurance, copay, etc.)
  Oth: Anything else (record in the notes)
Notes – Free text to elaborate on anything

Any time any of the observation fields change, you should start a new line. The time difference is then the duration for the prior line.

Figure B-2: Patient observation form, page 2

89
Appendix C

Revised Renewal Line Voicemail Script
Figure C-1: Revised Renewal Line Voicemail Script
Appendix D

Additional Time Study Observations

- All roles stay on task for the vast majority of the day. Items such as personal discussions, bathroom trips, coffee breaks, etc. were categorized as NV-Oth, which combines with several other items and so is not immediately visible in the quantitative data. However, all observers noted some surprise at the very low level of non-work activity compared to other organizations they have seen. There may be an element of bias, in that people being observed may alter their work habits due to the presence of the observer, but in no cases did the observers perceive any obvious examples of substantially altered workflow or behavior.

- The level of patient care not attributable to a visit (and hence not reimbursed under most current insurance plans) was surprisingly high. As IMA transitions to a PCMH model with capitated reimbursement this will become less of an issue but, during the transition period, the mismatch between resource expenditure and practice income may exacerbate discussions around appropriate panel sizes, billing, and provider compensation.

- The level of observed cross-pod interaction was very low. Most roles interact with other roles within their own pod several times a day but have very few, if any, interaction with IMA staff from other pods. Several current initiatives should help address this but the level of day-to-day knowledge sharing at the tactical level seemed very low.

- Each pod as its own “personality”, reflected in the cultural norms of each pod and
their perception of their own status relative to other pods. This is not inherently bad or good, and likely serves as a powerful way to build connections with an organization the size of IMA, but does present change implementation challenges as each pod views their own needs, workflows, and strengths differently from each other.

- In many cases, the observers saw many different members of the practice struggle with small IT problems that generated considerable rework and confusion: non-working printers, QuickSilver clinical messaging lockup, CRMS referral system's inability to save partial work, and inability for particular workstations to login toIDX scheduling. The team also observed many different deployments of the same function for different purposes (notably use of QuickSilver) and the use of different functions for the same purpose (email vs. QuickSilver vs. greensheets vs. mailboxes). This problem became more acute with cross-pod interactions or staff members floating between pods because of the lack of consistent use of tools.

- Huddle conduct was very inconsistent across pods and providers. Some teams did not conduct a huddle at all, some had a very perfunctory meeting with only some of the recommended roles present, and some had everyone and were very thorough. The location also varied, with some huddles conducted in the MA area “in the middle of things” with others conducted behind a closed door in the provider’s office. More thorough huddles appeared to correlate with pods and teams that the practice deemed as “higher performing” but this study did not determine whether this was a causative or correlative relationship.

- Many roles and pods have no structured plan for breaks, including lunch. In some cases where there is only one person present (e.g. nurses), the only observed solutions either disrupt the pod workflow or drive the staff member to eat lunch on the fly. Even in pods where there are multiple staff for the same role, the workload jump going from two to one staff is significant and resulted in large phone backlogs in at least two observed occasions.

- It is difficult to have conversations in the front office that are not audible to patients
at the desk talking to the coordinators. There were no observations of this causing any problem or noticeable patient discomfort but may present HIPAA concerns.

- A large portion of non-appointment patient contact takes place through the phone. Although some email and iHealthspace web messaging was observed for all roles, it was a small fraction relative to requests by phone. This was a source of consistent and unpredictable interruption for everyone observed. Several staff members with work history at other practices commented that IMA handled an unusually high volume of interaction via phone.

- There is low consistency of assignment of tasks to roles. The same task may be done by a coordinator, a medical assistant, or a nurse depending on the pod and provider. In several cases, there were locally designed “custom” workflows to support one particular provider. Although sometimes effective at the local level, this caused difficulty during staff substitutions.
Appendix E

Network Flow Model Screenshots
<table>
<thead>
<tr>
<th>Variable</th>
<th>Patient Population</th>
<th>Total Patients</th>
<th>On Meds</th>
<th>Rx Per Pt Per Year</th>
<th>Daily SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constants</td>
<td>Number Of Pods</td>
<td>Weeks Per Year</td>
<td>Days Per Year</td>
<td>WorkWeek</td>
<td>Work Days Per Week</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prescription Request Channels</th>
<th>% Rx By Apprk</th>
<th>% Rx By Web</th>
<th>% Rx By VMail</th>
<th>% Rx By Fax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prescription Request Channels</td>
<td>Appt Yield</td>
<td>Web Yield</td>
<td>VMail Yield</td>
<td>Fax Yield</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Excess Renewals</th>
<th>Excess Target</th>
<th>Chinnic and Stable Meds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target per year</td>
<td>Value of &quot;Excess Renewals&quot;</td>
<td>with a perfect system</td>
</tr>
<tr>
<td></td>
<td>Fraction of prescriptions that are chronic and stable</td>
<td>In the &quot;Excess Renewals&quot; calculation</td>
</tr>
<tr>
<td></td>
<td>Fraction of excess renewals that are caught by improved patient intervention</td>
<td>Assumption: Fraction of excess renewals that can be eliminated</td>
</tr>
<tr>
<td></td>
<td>Fraction of spilled renewals that could be done at the appointment with improved processes</td>
<td>Assumption: Fraction of spilled prescriptions that can be shift.</td>
</tr>
</tbody>
</table>

| Source            | OnCall rx log, 2010-2012 | OnCall rx log daily, excluding holidays/Sat/Sun | OnCall prescription log with processing against IDX log | Based on observed .125/.875 split for fax/vmail | Based on observed .125/.875 split for fax/vmail |

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Mean</th>
<th>I Uoh</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Source            | OnCall rx log, 2010-2012 | OnCall rx log daily, excluding holidays/Sat/Sun | OnCall prescription log with processing against IDX log | Based on observed .125/.875 split for fax/vmail | Based on observed .125/.875 split for fax/vmail |

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Mean</th>
<th>I Uoh</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Source            | OnCall rx log, 2010-2012 | OnCall rx log daily, excluding holidays/Sat/Sun | OnCall prescription log with processing against IDX log | Based on observed .125/.875 split for fax/vmail | Based on observed .125/.875 split for fax/vmail |

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Mean</th>
<th>I Uoh</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Source            | OnCall rx log, 2010-2012 | OnCall rx log daily, excluding holidays/Sat/Sun | OnCall prescription log with processing against IDX log | Based on observed .125/.875 split for fax/vmail | Based on observed .125/.875 split for fax/vmail |

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Mean</th>
<th>I Uoh</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Source            | OnCall rx log, 2010-2012 | OnCall rx log daily, excluding holidays/Sat/Sun | OnCall prescription log with processing against IDX log | Based on observed .125/.875 split for fax/vmail | Based on observed .125/.875 split for fax/vmail |

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Mean</th>
<th>I Uoh</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure E.2: Screenshot – Model inputs, page 2
## Prescription Output

- **Total Issued**: 183,474
- **Chronic Excess Rx**: 14,166
- **Spilled Rx (2-day)**: 7,782
- **Excess Rx Diverted**: 0
- **Spilled Rx Diverted**: 0

### Issued

- **Pre-Spill Spill Recovery Total**:
  - Central: 5,368
  - Web: 5,368

### Waste

- **Waste Pods**:
  - Central: 91,554
  - Total: 91,558

### Process Times

- **Average process time for a waste prescription** (weighted average of individual process times)
  - Central: 2 minutes, 54 seconds
  - Web: 2 minutes, 54 seconds

- **Variance**
  - All: 0
  - Per Day: 1.4
  - Per Pod: 0

### Central CoV

- **Central CoV**: 0.032

### Prescription Output

<table>
<thead>
<tr>
<th>COV</th>
<th>Web</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>

### Process Times

- **Total Fax**: 75,041
- **Total Web**: 53,214

### Process Recovery

- **Pre-Spill Spill Recovery Total**: 91,554

### Central

- **Central Total**: 91,554
- **Central CoV**: 0.032

### Variance

- **Variance**: All: 0
- **Variance**: Per Day: 1.4
- **Variance**: Per Pod: 0
Figure E-4: Screenshot – network adjacency matrix
Figure E-5: Screenshot – decision variables (current state)
Network Constraints

Out In Delta Sign Const. Justification

\begin{align*}
222882.1 & -7782 -215100 \leq 0 \\
\end{align*}

Source node

\begin{align*}
99336.09 & = 99336.09 \\
\end{align*}

Conservation

\begin{align*}
5697 & = 5697 \\
16442 & = 16442 \\
101407 & = 101407 \\
5697 & = 5697 \\
11559 & = 11559 \\
75041 & = 75041 \\
0 & \geq 0 \\
0 & = 183475 \\
0 & = 31625 \\
0 & = 31625 \\
\end{align*}

Sink node: complete prescriptions

\begin{align*}
0 & = 0 \\
0 & = 0 \\
0 & = 0 \\
0 & = 0 \\
0 & = 0 \\
\end{align*}

Sink node: redundant prescriptions

\begin{align*}
0 & = 0 \\
0 & = 0 \\
0 & = 0 \\
0 & = 0 \\
0 & = 0 \\
\end{align*}
<table>
<thead>
<tr>
<th>Path Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value Sign Const.</td>
</tr>
<tr>
<td>Match RxAtAppt to logs</td>
</tr>
<tr>
<td>Match RxFromWeb to logs</td>
</tr>
<tr>
<td>Match RxFromVmail to logs</td>
</tr>
<tr>
<td>Match RxFromFax to logs</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Central process - Web</th>
<th>Central process - Vmail</th>
<th>Central process - Fax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waste prescription from central process due to spillage</td>
<td>= 0</td>
<td>= 0</td>
</tr>
<tr>
<td>Prescription diverted due to excess reduction</td>
<td>= 0</td>
<td>= 0</td>
</tr>
<tr>
<td>Waste prescription from central process due to excess reduction</td>
<td>= 0</td>
<td>= 0</td>
</tr>
<tr>
<td>Prescription diverted due to excess reduction</td>
<td>= 0</td>
<td>= 0</td>
</tr>
</tbody>
</table>

Total Waste: 10,000
Total Prescription: 20,000
Total Prescription from Central Process: 10,000

Prescription diverted due to spillage: 0
Prescription diverted due to excess reduction: 0

Waste prescription from central process: 0

Central process: 0
Web: 0
Vmail: 0
Fax: 0

Note: Path constraints and values are placeholders for demonstration purposes.
Total Of Prescriptions

Total Patients w/ Prescriptions
Non-appointment Prescriptions
Appointment Prescriptions
Prescription Requests
Waste Requests
Excess Requests
Excess Renewal %
Spillover Prescriptions
Prescription Volume CoV

Per Pod
Central
All Practice

Daily Load
Per Pod
Central
All Practice

+1 SD 98%
Per Pod
Central
All Practice

+2 SD 98%
Per Pod
Central
All Practice

Should match to input, within rounding error, to verify model integrity

91921
Prescriptions issued outside appointments (not on a day with an arrived appointment)

91554
Prescriptions issued at appointments (on a day with an arrived appointment)

215100
Estimate of number of requests for prescriptions

31625
Estimate of number of requests that do not result in a prescription (duplicates, incomplete info, wrong patient, etc.)

14166
Estimate of number of requests that were not clinically necessary

7.7%
Estimate of fraction of prescriptions issued that were not clinically necessary

7782
Number of prescriptions that could have been handled at the appointment (2-day window) but were not

0.655
Standard Deviation/mean prescription demand per pod

0.000
Standard Deviation/mean prescription demand for centralized processing (depends on # of centralized pods)

0.175
Standard Deviation/mean prescription demand for full practice (equivalent to complete centralization)

25
Mean daily prescription request demand, each pod

0
Mean daily prescription request demand, central team

353
Mean daily prescription request demand, total practice

42
Mean + 1SD daily prescription request demand, each pod

414
Mean + 1SD daily prescription request demand, central team

58
Mean + 1SD daily prescription request demand, total practice

424
Mean + 2SD daily prescription request demand, each pod

476
Mean + 2SD daily prescription request demand, central team

58
Mean + 2SD daily prescription request demand, total practice
Average cycle time:

- **Average prescription cycle time:** hh:mm [including both processing and wait time]

- **Process redesign time saved:** 0

- **Specialization time saved:** 0

- **Total hours of processing time saved due to process redesign:**

- **Average cycle time reduction:** 0:00

- **Total FTE required for the complete practice (individual & centralized pods):** 4.07

- **Total FTE required for the complete practice on an average day:**
  - All Practice: 4.1
  - Per Pod: 0.3
  - Central: 0.0

- **Total FTE required for the complete practice on an average +1 SD demand day:**
  - All Practice: 4.8
  - Per Pod: 0.5
  - Central: 0.0

- **Total FTE required for the complete practice on an average +2 SD demand day:**
  - All Practice: 5.5
  - Per Pod: 0.7
  - Central: 0.0

By role:

- **PSC:** 0.0
- **MA:** 2.5
- **RN:** 1.3
- **MD/NP:** 0.0
- **Admin:** 0.0
- **Clinical:** 0.0
- **Prescriber:** 0.0

Average utilization per role:

- **PSC:** 1.7%
- **MA:** 18.2%
- **RN:** 9.1%
- **MD/NP:** 0.0%
- **Admin:** 0.0%
- **Clinical:** 0.0%
- **Prescriber:** 0.0%
Bibliography


Jane Hummer and Cristina Daccarett. Improvement in prescription renewal handling by application of the lean process, June 2009.


[38] Devon J. Price, School of Management Sloan, Institute of Technology Massachusetts, for Global Operations Program Leaders, and General Hospital Massachusetts. *Managing variability to improve quality, capacity and cost in the perioperative process at Massachusetts General Hospital*. 2011.


