Information-Enabled Decision-Making in Health Care: EHR-Enabled Standardization, Physician Profiling and Medical Home

by

Lori Rebecca Pelletier

A Dissertation

Submitted to the Graduate Faculty of

Worcester Polytechnic Institute

in partial fulfillment of the requirements for the

Degree of Doctor of Philosophy

in

Manufacturing Engineering

April 20, 2010

APPROVED:

Sharon Johnson, PhD, Associate Professor of Management, WPI – Advisor

Isa Bar-On, PhD, Professor of Mechanical Engineering, WPI

Diane Strong, PhD, Professor of Management, WPI

Edward Westrick, MD, PhD, Vice President of Medical Management, UMMHC

Qi Zhou, MD, MBA, MS, and CPH, Director of Quality Management and Health Informatics, Tufts Health Plan

Abstract

Health care today harms too frequently and routinely fails to deliver its potential benefits. Significant evidence suggests that high quality primary care can positively affect health outcomes. I explored three related topics mentioned frequently in current United States health reform – Electronic Health Records (EHR), physician profiling and Medical Home. An investment in these areas is expected to significantly improve quality of care and efficiency; however, there is only a patchwork of evidence supporting such claims.

To achieve ***EHR*** promises, my research employed a standardization lens to study the dynamics between EHR embedded structures and primary care processes. Using grounded theory, a standardization dynamics model was created describing the influencers, conditions and consequences of the process state. A matrix of two conditions, information exchange and patient complexity, identified four distinct pathways that require a different balance between standardization and flexibility. The value of such pathways is that they frame choices about how to use embedded IT structures to support effective delivery processes.

Physician profiling is an emerging methodology used in health care quality improvement programs. Efforts to measure performance at the individual physician level face a number of challenges, including the need for sufficient sample size to support reliable measurement. A process for creating a physician profiling model was developed, and a model designed for a case study site. Results indicate that reliable physician profiling is possible across care domains using a hierarchical composite model.

Patient-Centered Medical Home (PCMH) is a new care delivery approach for providing comprehensive primary care that seeks to strengthen the physician-patient relationship. This exploratory study utilizes Pearson correlation coefficients to test four hypotheses about

relationships between two sources of data: (1) PPC-PCMH Survey results that measure adoption of PCMH structures and (2) patient experience data from Massachusetts Health Quality Partners (MHQP). The results showed that the PPC-PCMH structures of access and communication were negatively correlated with the related patient experience measure.

This study contributes to the literature by addressing deficiencies in how EHR-enabled processes, physician profiling models and Medical Home constructs are measured, to support improved outcomes.

Acknowledgements

I would like to thank my advisor, Professor Sharon Johnson, for an amazing four years. Her high level of expectations and her kind, humble, and encouraging approach are qualities every student, manager, teacher and parent should have. I thank her for her patience, guidance, support, timely feedback and most of all for her unprecedented dedication to my journey - I am forever grateful!

To my dissertation committee, Professor Diane Strong, Professor Isa Bar-On, Dr. Edward Westrick and Dr. Qi Zhou, my sincere gratitude for their active participation and their kind support.

Special thanks to Professor Diane Strong for sharing her expertise and providing guidance on qualitative analysis, paper writing and theory building skills. And, special thanks to Professor Isa Bar-On for always encouraging me to think "big" in both my writing and framing my contributions. They both have given me an endless supply of guidance and support in the past four years that have not only shaped my current research but also my ability as a future researcher.

Special thanks to Dr. Edward Westrick for providing me the opportunity to work at UMMHC as a research assistant and conduct state-of-the-art research. I am very grateful for all of his long hours and patience in introducing me to health care operations, health care research and health policy. Many thanks to Elaine Fontaine, Dr. Alan Krinsky, Christine Cernak, Jeanne McBride and Dr. Robert Klugman for welcoming me into the UMMHC community and their guidance, expertise, and hard work on our research topics.

Table of Contents

List of Tables

List of Figures

0 Introduction

Health care (HC) delivery in the United States (US) is undergoing dramatic changes, driven by two forces: (1) national attention because of rising costs, increasing demand for HC services as the baby boomers become elderly, and increasing evidence of quality problems in HC delivery (Bates et al., 2001; IOM, 2001), and (2) the promise of information technology (IT) designed for the HC context, commonly called HIT (Helm and Hansen, 2004). In 2007, the US spent $2.2 trillion on health care or 16% Gross Domestic Product (CMS, 2007). At the same time, US health care delivery ranked last out of sixteen countries based on measures of life expectancy, mortality due to cancer, circulatory diseases, respiratory diseases, mental disorders, as well as infant mortality and self-reported health status (World Health Statistics, 2009). The Institute of Medicine has recognized the need to improve quality of care in the United States, documenting that estimated errors and poor processes cause between 44,000 and 98,000 deaths and one million injuries annually (IOM, 2001). This estimate compares with 43,664 deaths a year from motor vehicle accidents and 30,896 deaths associated with firearms including gang warfare, self defense shootings and criminals killed by police (Heron et al., 2009). The Committee on Engineering and the Health Care System (which was supported by funding agencies including NSF and NIH, as well as the National Academy of Engineering and the Institute of Medicine) suggested that a partnership between engineers and health care professionals is needed to develop an effective delivery system (Lawrence, 2005). Significant opportunity exists to apply the process design, quality, and information technology tools that have been used in other manufacturing and service industries to dramatically improve performance.

In March 2010, the Patient Protection and Affordable Care Act was signed into law, with the goal of boosting access to health insurance for all Americans, reducing the escalating growth of health care costs and delivering the quality of care that the US health care system is capable of delivering (Kaiser Family Foundation, 2009). This dissertation focuses on developing information-enabled decision making methods to support improvement in quality and health system performance. The areas considered are standardization with Electronic Health Records (EHR), physician profiling, and Patient-Centered Medical Home, three topics repeatedly mentioned in health reform proposals from six committees (Kaiser Family Foundation, 2009).

The dissertation is focused on primary care because the majority of medical care is delivered in a primary care setting (Green et al., 2006) and significant evidence suggests that high quality primary care can positively affect health outcomes (Bodenheimer et al., 2002; Bodenheimer, 2006; Starfield, 2000; Starfield et al., 2005). In 2006, approximately 900 million visits were made to office-based physicians in the United States (Cherry et al., 2008), about 60% of which were visits to primary care providers. Fifty percent of these were made by patients with one or more chronic conditions. At the same time, a shortage of primary care physicians has added a sense of urgency to accelerate change. It is estimated that primary care physicians must work 18 hours every day, given the current state of the United States health care (HC) system, just to provide all of the recommended preventive and chronic care services that patients need. Addressing acute problems would take even more time. As a result, only half of the appropriate care is actually being provided to patients (McGlynn et al., 2003). In addition, most errors in primary care are preventable (Fischer et al., 1997), rendering primary care ripe for improvement.

My first dissertation topic explores a process improvement enabler, *standardization,* by examining the dynamics between the embedded structure of an *Electronic Health Records*

(EHR) system and the data, tasks and roles in a primary care process. Standardization has played a significant role in improvement methodologies like lean and Six Sigma, by reducing unwanted variation and sharing best practices (Parks, 2003; Sharrock, 2007; Wood, 2004). Some authors advocate that a process cannot be improved until it is standardized (Imai, 1986; Liker, 2004). Electronic Health Record (EHR) systems hold the promise of significantly improving care delivery by maintaining patient information and medication lists, storing medical notes related to patient encounters, and providing tools for ordering prescriptions and tests. However, as implemented, EHRs have not been associated with better quality ambulatory care (Linder et al., 2007). Recent literature studying outcomes after an EHR implementation suggests focusing on health care processes as a means to deliver on EHR promises (Bates et al., 2003; Hing et al., 2009; Poissant et al., 2005; Zhou et al., 2009). The complexity and variation in health care requires adaptable processes, creating a dynamic tension between standardization and flexibility. We use a process lens to examine standardization patterns in primary care clinics, considering the EHR embedded structure as an intervening condition.

Our analysis is based on a grounded theory approach, using data collected as part of a broader study (NSF SES-0826842) in which 112 interviews were conducted by WPI professors with primary care physicians, their medical staff and medical management in a large multi-specialty practice before, immediately after and over one year after an EHR implementation. The work contributes to an understanding of a specific enabler of process improvement, standardization, and the opportunity to leverage the structure of EHR systems to support it. In a reform proposal submitted to Congress, Baker et al (2009) recommend an investment in "meaningful and effective use of HIT". The findings of this study have the potential to support

implementation planning for the 70% of office-based physicians who are not yet using an EHR system along with those who are seeking to more effectively use their EHR system.

Physician Profiling, my second dissertation topic, is gaining increasing attention as health care organizations and insurance companies attempt to find incentives to improve quality of care and efficiency. Health plans are currently entertaining a pay-for-performance contract structure with the hope of identifying and rewarding those physicians associated with higher quality of care and efficiency. Primary care physicians tend to disregard such report cards because single payer claims are not representative of individual performance and often use specialty focused measures (e.g., Diabetes measures in pediatrics) leading to small sample sizes and poor reliability (Scholle et al., 2009). Although pay-for-performance is considered better than fee-for-service (Draper, 2009), some fear that pay-for-performance will result in pay-for-compliance instead of pay-of-excellence (Berwick, 2009). The objective of this study was to create a model to reliably differentiate primary care physicians across quality of care and efficiency measures using data from multiple payers and data sources. First, a process for creating a physician profiling model is presented, which incorporates key value criteria from the literature. The process is then used to create a profiling model at a case study site using reliability calculations suggested by Adams (2009). The structure of this flexible, expandable and reliable physician profiling model could also support the Senate Finance Committee America's Healthy Future Act of 2009 (2009) call for "moving from pay-for-reporting to pay-for-performance based on measures reflecting overall quality and coordination of care".

My third dissertation topic explores the relationships of *Patient-Centered Medical Home (PCMH)* and measures of patient experience. Medical Home is a new care delivery model that seeks to strengthen the physician-patient relationship by replacing episodic care based on

illnesses and patient complaints with coordinated care and a long-term healing relationship. The PCMH model is a list of six principles recognized by physician associations and payers as a means of achieving the Institute of Medicine's vision of care that is safe, equitable, efficient, effectively, patient-centered and timely (IOM, 2001; Rosenthal, 2008). The Physician Practice Connections Patient Centered Medical Home survey (PPC-PCMH) can be used by practices to evaluate their progress toward specific structural characteristics associated with the PCMH (www.ncqa.org , accessed May 28, 2009). These characteristics include, for example, an ongoing relationship with a personal physician and use of information technologies to support coordination of care.

While the PCMH model is gaining support, there is still very little evidence linking the model with better outcomes (Rosenthal, 2008; Reid et al., 2009). This exploratory study utilizes Pearson correlation coefficients to test four hypotheses about relationships between two sources of data, (1) the PPC-PCMH Survey results that measure adoption of PCMH structures and (2) patient experience data from Massachusetts Health Quality Partners (MHQP) to explore outcomes. Our analysis is based on the results gathered for 16 practices of a multi-clinic health care organization. This study provides a foundation for studying additional relationships between the structures assessed by the PPC-PCMH instrument and process outcomes, including patient satisfaction and quality measures. This study facilitates information-enabled decision making in order to gain "support for new models of delivering care through medical homes" (page 3 of The Obama Plan: Stability & Security for all Americans, 2009).

In conclusion, the goal of this dissertation is to provide empirical evidence and methodologies to support information-enabled decision making regarding three related topics in a primary care setting; EHR-enabled standardization, physician profiling, and Medical Home.

While these topics are repeatedly mentioned in 2010 health care reform proposals as solutions to the US health care crises, there is a patchwork of evidence to support such claims. This study uses qualitative and quantitative methods to analyze empirical data in two different health care organizations. The models and analyses developed can be used to support decision-making that leads to better health care delivery.

Thesis Organization

The remaining chapters of this thesis are organized according to my three dissertation topics. EHR-enabled standardization is discussed in Chapter 1. Physician profiling is discussed in Chapter 2 with detailed calculations and future studies presented in Chapter 3. Patient-Centered Medical Home is discussed in Chapter 4. A summary of the thesis conclusions, including a list of presentations and awards, is presented in Chapter 5, with recommendations for future research.

References

Adams, J. L. "The Reliability of Provider Profiling: A Tutorial," *Rand,* 2009.

Baker, H., T. Daschle, and B. Dole. "Crossing our Lines: Working Together to Reform the U.S. Health System," June 17, 2009.

Bates, D. W., M. Cohen, L. L. Leape, J. M. Overhage, M. M. Shabot, and T. Sheridan. "Reducing the Frequency of Errors in Medicine using Information Technology," *Journal of the American Medical Informatics Association : JAMIA* (8:4), Jul-Aug, 2001, pp. 299-308.

Bates, D. W., M. Ebell, E. Gotlieb, J. Zapp, and H. C. Mullins. "A Proposal for Electronic Medical Records in U.S. Primary Care," *Journal of the American Medical Informatics Association* (10:1), January 1, 2003, pp. 1-10.

Berwick, D. M., and S. King. "Web Institute of Health Improvement: Road to a Better Health Care System," September 24, 2009. Available from: http://www.ihi.org/IHI/Programs/AudioAndWebPrograms/WIHI.htm.

Berwick, D. M. "A User's Manual for the IOM's 'Quality Chasm' Report," *Health Affairs* (21:3), May 1, 2002, pp. 80-90.

Bodenheimer, T. "Primary Care--Will it Survive?" *The New England Journal of Medicine* (355:9), August 31, 2006, August 31, 2006, pp. 861-864.

Bodenheimer, T., E. H. Wagner, and K. Grumbach. "Improving Primary Care for Patients with Chronic Illness," *JAMA : The Journal of the American Medical Association* (288:14), Oct 9, 2002, pp. 1775-1779.

Centers for Medicare and Medicaid Services (CMS). "National Health Expenditures: 2007 Highlights," Washington, D.C.: 2009.

Cherry, D. K., E. Hing, D. A. Woodwell, and E. A. Rechtsteiner. "National Ambulatory Medical Care Survey: 2006 Summary," *National Health Statistics Report* (3), Aug 6, 2008, pp. 1-39.

Draper, D. A. "Physician Performance Measurement: A Key to Higher Quality and Lower Cost Growth Or a Lost Opportunity?" *Health Systems Change Brief* (3): June, 2009.

Green, C. J., P. Fortin, M. Maclure, A. Macgregor, and S. Robinson. "Information System Support as a Critical Success Factor for Chronic Disease Management: Necessary but Not Sufficient," *Int J Med Inform.* (75:12), Aug 16, 2006, pp. 818-828.

Helm, S.,2nd, and H. C. Hansen. "Information Technology in the Interventional Pain Practice: Electronic Medical Records, Practice Management Software, and Document Management," *Pain Physician* (7:3), Jul, 2004, pp. 357-364.

Heron, M., D. Hoyert, S. Murphy, J. Xu, K. Kochanek, and B. Tejada-Vera. "Deaths: Final Data for 2006," *National Vital Statistics Report* (57:14), 2009.

Hing, E. S., C. W. Burt, and D. A. Woodwell. "Electronic Medical Record use by Office-Based Physicians and their Practices: United States, 2006," *Advance Data* (393), Oct 26, 2007, pp. 1-7.

Imai, M. *Kaizen (Ky'Zen), the Key to Japan's Competitive Success*, Random House Business Division, New York, 1986.

Institute of Medicine (U.S.). Committee on Quality of Health Care in America. *Crossing the Quality Chasm : A New Health System for the 21st Century*, National Academy Press, Washington, D.C., 2001.

Kaiser Family Foundation. "Side-by-Side Comparison of Major Health Care Reform Proposals," *Focus on Health Reform,* September 18, 2009. Available from: www.kff.org. Accessed October 20, 2009.

Lawrence, D. *"Bridging the Quality Chasm," Building a Better Delivery System: A New Engineering/Health Care Partnership,* P. P. Reid, W. D. Compton, J. H. Grossman and G. Fanjiang (eds.), National Academy Press., Washington D.C., 2005.

Liker, J. K. *The Toyota Way: 14 Management Principles from the World's Greatest Manufacturer*, McGraw-Hill, New York, 2004.

Linder, J. A., J. Ma, D. W. Bates, B. Middleton, and R. S. Stafford. "Electronic Health Record use and the Quality of Ambulatory Care in the United States." *Archives of Internal Medicine* (167:13), 2007, pp. 1400-1405.

McGlynn, E. A., S. M. Asch, J. Adams, J. Keesey, J. Hicks, A. DeCristofaro, and E. A. Kerr. "The Quality of Health Care Delivered to Adults in the United States," *The New England Journal of Medicine* (348:26), Jun 26, 2003, pp. 2635-2645.

Obama, B. "the Obama Plan: Stability & Security for all Americans," February 26, 2009.

Parks, C. M. "The Bare Necessities of LEAN," *Industrial Engineer: IE* (35:8), 08, 2003, pp. 39.

Poissant, L., J. Pereira, R. Tamblyn, and Y. Kawasumi. "The Impact of Electronic Health Records on Time Efficiency of Physicians and Nurses: A Systematic Review," *Journal of the American Medical Informatics Association: JAMIA* (12:5), Sep-Oct, 2005, pp. 505-516.

Reid, R. J., P. A. Fishman, O. Yu, T. R. Ross, J. T. Tufano, M. P. Soman, and E. B. Larson. "Patient-Centered Medical Home Demonstration: A Prospective, Quasi-Experimental, before and After Evaluation," *The American Journal of Managed Care* (15:9), Sep 1, 2009, pp. e71-87.

Rosenthal, T. C. "The Medical Home: Growing Evidence to Support a New Approach to Primary Care." *Journal of the American Board of Family Medicine (Formerly Journal of the American Board of Family Practice)* (21:5), 2008, pp. 427.

Scholle, S. H., J. Roski, D. L. Dunn, J. L. Adams, D. P. Dugan, L. G. Pawlson, and E. A. Kerr. "Availability of Data for Measuring Physician Quality Performance," *The American Journal of Managed Care* (15:1), Jan, 2009, pp. 67-72.

Sharrock, R. "Road to Gold," *Industrial Engineer: IE* (39:5), 05, 2007, pp. 44-48.

Starfield, B. "Is US Health really the Best in the World?" *Journal of the American Medical Association* (284:4), July 26, 2000, July 26, 2000, pp. 483-485.

Starfield, B., L. Shi, and J. Macinko. "Contribution of Primary Care to Health Systems and Health," *Milbank Quarterly* (83:3), 09, 2005, pp. 457-502.

Wood, N. "Lean Thinking: What it is and what it Isn't," *Management Services* (48:2), 02, 2004, pp. 8-10.

World health organization. "World Health Statistics Report," 2009.

Zhou, L., C. S. Soran, C. A. Jenter, L. A. Volk, E. J. Orav, D. W. Bates, and S. R. Simon. "The Relationship between Electronic Health Record use and Quality of Care Over Time," *Journal of the American Medical Informatics Association* (16:4), July/August, 2009, pp. 457-464.

1.0 Employing a Process Lens to Define Pathways in Primary Care

1.1 Introduction

Primary care physicians provide acute, chronic, and preventive care to a diverse population in an outpatient setting, including a growing number of patients with multiple chronic diseases and comorbid conditions. At the same time, scientific advances and new technology constantly change the practice of medicine, so primary care physicians must integrate and manage an enormous amount of information and biomedical knowledge while delivering patient-centered care. In this paper, we examine standardization, supported by information technology, as a means to advance quality, efficiency and satisfaction within primary care microsystems. Microsystems are the small, functional frontline units that provide most health care to most people (Nelson et al., 2007).

In the context of processes, standardization refers to creating uniform tasks and sequences for carrying out tasks, eliminating variation and unnecessary work. Lean and six sigma methods include standardization as a critical component of process redesign; such methods have resulted in significant performance improvement in manufacturing (Imai, 1986; Liker, 2004; Parks, 2003; Sharrock, 2007; Wood, 2004). Yet in the context of both variable and knowledge-intensive work, finding a good balance between standardization and flexibility can be critical to improving performance and competitive advantage (Brown and Duguid, 2000; Bruce et al., 2007; Stalk et al., 1992). Our study examines the balance between standardization and flexibility in primary care delivery processes, recognizing that individual provider style and the variability and complexity of patient conditions and treatment regimens are innate system characteristics.

Electronic health record (EHR) systems provide structure to support more efficient health care delivery processes and better quality care, by maintaining patient information such as medical history and medication lists, creating and storing medical notes related to each patient

encounter, decision support including reminders and interaction alerts, and tools for ordering prescriptions and tests. Such systems structure data and processes, and the resulting standardization has created significant advantages for companies implementing other types enterprise systems (Davenport et al., 2004; Volkoff et al., 2005). However, to date, increases in quality of care and efficiency have not been realized with EHR systems (Linder et al., 2007; Makoul et al. 2001; Pizziferri et al., 2004). In addition, current implementations are likely to involve vendor-built EHR systems, where embedded structures must be integrated with care delivery processes by exploiting customization tools and redesigning workflow (Maxwell, 1999; Poulymenopoulous et al., 2003). While the importance of combining IT implementation and process redesign is commonly understood, health care organizations often wait to make process design changes until after the implementation, which has been associated with inefficiencies in regaining capacity and EHR aborted implementations (Scott et al., 2005). Recent literature suggests that more understanding of the dynamic interface between EHRs and health care processes is needed (Hing et al., 2009; Zhou et al., 2009).

With the growing complexity of primary care, the challenge is how to integrate the EHR embedded structure with the processes of primary care, which are designed to meet the individual needs of a diverse patient population. Current improvement models offer conceptual frameworks for using IT systems to improve quality, safety, and IT adoption in the health care context (Nemeth et al., 2008; Carayon et al., 2006; Wagner, 1998), there are currently very few models at the microsystem level that offer practical guidelines as to how to integrate IT and process design to achieve better outcomes (Gittell et al., 2009). We address this gap in the literature by using a process lens to study how, why and with what consequences standardization and flexibility coexist in a primary care process using an EHR embedded structure.

We use a grounded theory approach, based on 78 interviews of primary care physicians and medical staff, before and after implementation of an EHR, at a multi-site, multi-specialty practice we call "Medical Clinic". Grounded theory supports the discovery of concepts and relationships in raw data and a process for organizing these into a theoretical explanatory scheme (Strauss and Corbin, 1998). Through longitudinal interviews and a constant review of the literature, a robust foundation for theory development regarding the dynamics between EHR structures and primary care delivery processes is possible.

We develop a dynamic model that defines the process state of standardization, characterizing dimensions that influence the extent and type of standardization in a delivery process. We also discovered distinct pathways of the primary care process that require different levels of flexibility and standardization. The value of such pathways is that they identify and frame choices about how to use embedded EHR structures to support effective care plan delivery, going beyond generic advice to redesign processes in conjunction with IT implementation. Recognizing and designing for specific pathways before an EHR implementation can minimize the tension from a one-size fits all approach and can likely increase its meaningful use. A process design tool was also created to qualitatively represent current states and to plan for the future states of standardization across the organization and for each pathway. The contribution is a grounded theory that links process improvement decisions related to standardization with the opportunities created by embedded structures in the EHR system. The findings of this study have the potential to support the implementation planning of the 70% of office-based physicians who are not yet using a full or partial EHR system (Hing et al., 2007), along with those who are implementing improvements after EHR adoption.

1.2 Literature Review

The literature review explores five components deemed critical for understanding the opportunity of standardization in primary care. The critical components are the primary care environment, the role standardization has played in process improvement, health information technology, EHR effects on outcomes, and broad models for improvement. We discuss each of these in turn.

1.2.1 Information and the Primary Care Environment

Primary care physicians are challenged with more to know, more to manage, more to watch, more to do, and more people involved in doing it than at any time in the nation's history (IOM, 2001). This challenge is fueled by the advancements in medical science and technology, which have grown faster than the systems required to deliver them safely, effectively and efficiently, and patient preferences to use such technology. The sudden surplus of information in health care has been associated with misinterpretations, lost information or incomplete information leading to unintended consequences at the patient level (IOM, 2001; etc). For example, medication errors have been tightly linked to poor information exchange (e.g., accuracy and completeness), information availability (e.g., information pertaining to existing medications and interactions) and task completeness (e.g., patient getting the right prescription). Fischer et al. (1997) found that 83 percent of all errors in eight primary care practices were preventable. While some of these consequences can be fatal, many cause non-value added work consuming resources and time.

Information exchange has been studied to evaluate the effectiveness of patient-physician communication (Apter et al., 2008) and the data quality and availability of information exchanged between health care entities (Barua et al., 2007; Frieling, 2009). Information exchange (IE) as defined in this paper is the information related to patient care exchanged

between roles during a primary care visit.   Figure 1.1 illustrates a typical flow of patient

information between roles (upper row of ellipses) and tasks (second row of ellipses), starting on

the left when a patient arrives at the clinic. The patients' medical record and other sources (e.g.,

faxes with lab results), as well as contact with the patient during the visit, provide information

about the patients' history and condition to support the visit. For example, for an acute visit,

patient information may be gathered prior to the visit through a phone triage carried out by a

nurse.
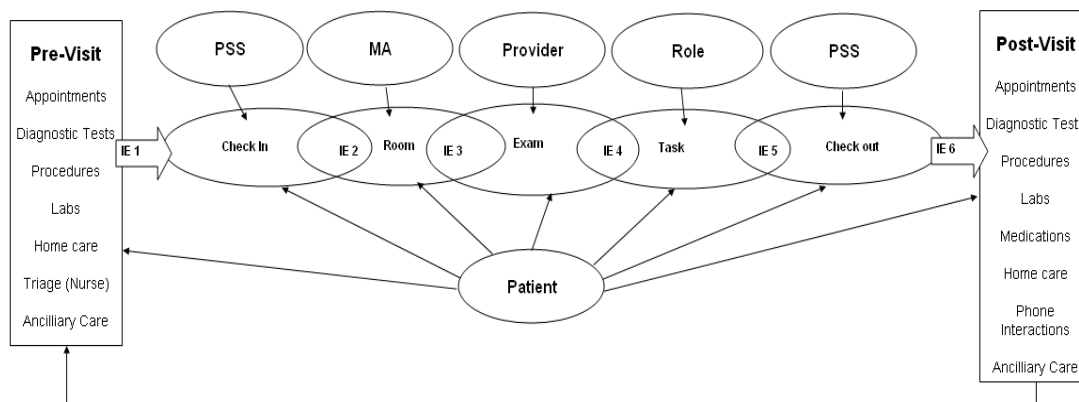
Figure 1.1 Process of Primary Care Visit

Upon check-in additional information is exchanged between the patient and the patient

support specialist (PSS), denoted as IE1 in Figure 1.1. The PSS then notifies the medical

assistant (MA) as to the patient arrival (IE2) along with any additional information (e.g., what

the patient is wearing for a personal greeting by MA). The MA then rooms the patient, obtains

and documents vitals, restates the general purpose of the visit and performs any pertinent

protocols (e.g., strep test, nebulizer). The physician then uses the patient information (IE3)

during the exam and through interactions with the patient determines a diagnosis or health status

and care plan for the patient. The physician then documents and communicates this information

(IE4) in the encounter and with the patient (e.g., patient education, forms, medications, orders,

referrals, future appointments). At this point, the physician might request additional station

medications, supplies or tests from a nurse prior to finishing the exam. The care plan and visit

summary is then delivered to the PSS (IE5) at checkout for post-visit needs shown in Figure 1.1.

Primary care physicians have grown to rely on some form of health information

technology as a means to manage the plethora of information exchanges during routine and

knowledge intensive patient visits. In this study, we explore information exchanges during

primary care visits to understand under what conditions information exchanges can be

standardized and those which flexibility is needed.

1.2.2 Standardization

Process standardization has been pursued in nearly every industry, generating remarkable

gains in quality and efficiency in many, by reducing unwanted variation and eliminating non-

valued added activities (Imai, 1986; Liker, 2004; Parks, 2003; Sharrock, 2007; Wood, 2004).

Process standardization is not new to health care. Standardization methodologies like

Standardize-Do-Check-Act (Nelson et al., 2007) and, more recently, lean (Dickson et al., 2009)

have been widely adopted.   At any point in time, processes have some level or extent of

standardization; process improvements or information systems may increase the extent or

standardization (or conversely, limit its extent by allowing more flexibility). In a medical clinic,

some level of process standardization exists prior to an EHR implementation. For example,

paper charts might be delivered to clinics the day before the scheduled visits and staff collects

recent history in preparation for the exam room visit. Providers are handicapped without such

information to properly diagnosis the problem and provide the best treatment plan. When

information is missing, potential delays exist in the exam room visit and hence, providers wait.

If the same process is followed for each patient, the process begins with a greater extent of

standardization prior to such information being made available via an EHR. Standardizing processes is linked to capacity in that if the provider can efficiently and effectively communicate a diagnosis and treatment plan to the patient during the visit, the provider can move on to the next patient.   However, too much standardization in primary care can stifle the ability for physicians to capture the essence of the patient complaint, diagnosis and care plan and is bounded by the need for flexibility.

Patients, the inputs to HC processes, exhibit much more variability and complexity than most manufacturing inputs. The manufacturing sector can sometimes control quality by reducing the variability in input quality, an avenue not typically open to HC practices. Patient input variability refers to the type and range of patient conditions, rather than an alternative usage that focuses on the uncertainty about the condition of an individual patient, which we refer to as complexity (Argote, 1982). In addition, physicians are highly skilled professionals whose work cannot be completely standardized.

However, too much standardization could stifle innovation in an environment where customers value output variability or unique products to meet their needs and too much flexibility could hinder consistency where the customers value a narrow range of products or services. Finding a good balance between standardization and flexibility is a management challenge. Hall and Johnson (2009) define the concept of artistic processes, which support flexibility and creativity in changeable environments requiring judgment-based work and when customers value variation in the end product. Standardization can be used before and after 'artistic' process activities, to ensure they operate effectively. Primary care physicians typically provide customized care plans based on the health care needs and personal preferences of the patient. The difficulty in standardizing health care processes lies in transforming the patient

complaint and patient data to a meaningful, effective treatment plan with, in some cases choices

for the patient. The need for flexibility is driven by the need to manage complexity, with

potentially incomplete and conflicting data and a vast array of possible diagnoses, rather than the

need to support creativity. As with artistic processes, peripheral tasks can be standardized to

alleviate unwanted variation, promote consistency around the physician's work and provide flow

of the process to support capacity. However, the transformation of patient knowledge in a

primary care setting is not intended to yield new or artistic ways of treating a diagnosis.

Mass customization is another approach that has been used to balance standardization and

flexibility (Hall and Johnson, 2009). In mass customization, standardized processes produce

products using modular elements with limited consumer choices, similar to the automotive

industry. A similar approach might work for healthy patients who have routine complaints

requiring well-known treatment plans that might reflect patient preferences.

A method used in lean manufacturing to support standardization is to identify value

streams and create processes focused around these streams. A value stream is a group of

dedicated resources and tasks that collectively deliver a service or product (Womack and Jones,

2003). Value streams are identified by organizations and studied using lean methodologies to

reduce unwanted waste and provide seamless flow of the product or service. HC processes that

represent particular care delivery processes, e.g., care teams for diabetes have reduced variation

and complexity by organizing according to value streams (Bodenheimer et al., 2002;Dickson et

al., 2009; Litvak et al., 2005), but only a few medical conditions are common enough to devote

specific resources to and/or have clear care guidelines. Current EHR implementations are likely

to involve vendor-built EHR systems, which have embedded structure related to tasks and their

sequence, which must be integrated with the complexity and workflow of care delivery processes

(Pizziferri et. al., 2005). EHR systems may constrain choices related to value stream planning. If is also difficult to identify value streams broadly (for healthy and more complex patients, for example) because what appears to be a straightforward visit turns out to be complex. However, similar pathways form during the exam room as primary care physicians decide the diagnosis and care plan for patients of varying complexity. In this study, we explore the feasibility of defining pathways in primary care integrated with the EHR embedded structures as a mechanism to standardize those processes that can be standardized and allow for flexibility those processes that need more flexibility.

1.2.3 Health Information Technology

Early HIT provided administrative support to hospitals, similar to systems supporting commercial organizations (Berner et al., 2005). Subsequently, clinical decision support systems (Kaplan, 2001), as well as single-purpose applications to support labs and pharmacies (Berner et al., 2005) were created. Multi-media technology supported IT for telemedicine (Goldschmidt, 2005). More recently, computer-based physician order-entry (CPOE) and electronic health records (EHR) have been introduced. CPOEs and EHRs, when tied with lab and pharmacy systems, create a comprehensive EHR. An EHR is comprehensive if it includes four features: computerized orders for prescriptions, computerized orders for tests, electronic reporting of test results, and storage of clinical notes (Hing et al., 2007).

Physicians in the U.S. have long resisted HIT in care delivery (as compared to billing and financial management), e.g., only 17% of U.S. physicians as of 2001 used electronic health records (Bates et al., 2003). These usage rates contrast with a 90% rate in Sweden, 88% in the Netherlands, 58% in the United Kingdom, and over 95% in Israel (Poon et al., 2006). There is a growing belief in the U.S. that applying IT and other process design tools can dramatically

increase the efficiency and quality of HC delivery processes (Reid et al., 2005). As a result, insurance companies are changing re-imbursement policies as incentives for adopting HIT and federal agencies are providing incentives for such systems. Physician groups are responding, evidenced by the 22% and 60% increase in EHR use by physician offices since 2005 and 2001 respectively (Hing et al., 2007). With the change in incentives, the increased capabilities of current HIT and EHR, and the growing realization of the potential value of HIT to HC delivery, physician resistance to HIT is no longer the key issue. The key issue is how to revise HC delivery processes to use these systems to significantly improve HC quality and efficiency (Ovretveit et al., 2007).

Volkoff et al. (2007) discuss IT embedded elements of an enterprise system (ES) as routines, roles and data that can structure process tasks; such systems have properties comparable to that of an EHR system. This study focuses on primary care processes with the IT embedded elements serving as an intervening condition. That is, while an IT artifact can dictate the sequence of tasks for transactions and documentation within the IT system, other peripheral tasks in primary care occur outside of the system. One limitation of migrating properties of ES systems to EHR systems relates to variability. ES systems limit the number of choices in an embedded task to avoid excess complexity to support standardization, which is an expected outcome of implementing an ES (Volkoff et al., 2007). In health care, accommodating patient and treatment regime variability in completing tasks is an important ingredient in delivering patient-centered care. In this paper we examine how process and IT design decisions can embrace the complexity of primary care delivery in the presence of EHR embedded structures.

1.2.4 EHR Effects on Outcomes

Studies have shown mixed results on the effects of EHR systems on quality of care, efficiency and satisfaction outcomes. The results of a study evaluating physician and nurse satisfaction of an emergency room EHR indicated that physicians and nurses were generally satisfied with the EHR but did not attribute it to better patient care and worried about confidentiality and time to use such a system (Likourezos et al., 2004). In the context of patient and provider satisfaction, role changes brought about by the EHR may have unexpected consequences (Davidson and Chismar, 1999). With substitutability within and between roles, patients may be less satisfied when they cannot see their primary care provider and HC providers may feel the loss of professional and social identity (Davidson and Chismar, 2007; Lamb and Kling, 2003). Even practices that have had an increase in patient satisfaction and productivity from the system "go live" recognize that there are larger gains to be realized by refining the workflows (Maxwell, 1999).

Studies of EHR effects have shown small or no increases in physician productivity (Pizziferri et al., 2005; Bates et al., 1994), and quality of care (Linder et al., 2007; Zhou et al., 2009). Results from Linder et al. (2007) and Zhou et al. (2009) suggest that simply implementing EHRs is unlikely to result in improved quality. The quality measures used in these studies may have been far removed from the delivery process and not sensitive enough to reveal improvements in quality of care attributable to the adoption and use of EHR. A closer focus on the process of delivering care could reveal improved indicators of quality performance. For instance, Yackel and Embi (2010) found a new set of errors imposed by the implementation of a commercial EHR in reporting of test results. Erroneous test results are problematic for a patient's care plan and represent one of the biggest safety issues in outpatient care (Poon et al.,

2004). Using quality measures like Hemoglobin A1C control as a process improvement gauge

does not include steps in the methodology to identify test result management errors.

1.2.5 High-Level Models for Improvement

One reason that studies have shown only limited effects on outcomes may be because

delivery processes and EHR functionalities must be designed in tandem. Some studies suggest

focusing on better use of the IT artifact, e.g., intensifying the use of key EHR features, such as

clinical decision support (Zhou et al., 2009) and quality reporting and registries (Linder et al.,

2007) as the solution. Other studies suggest focusing on organizational artifacts, e.g.,

intensifying clinician skills to enhance clinician-patient communication (Frankel et al., 2005) and

designing workflows (Murray, 2003).

Socio-technical literature suggested models that embrace both IT and organizational

change levers suggesting dynamics between the people, technology, environment, and

organization are unavoidable and essential in understanding performance measures and medical

staff outcomes (Karsh, 2004; Carayon et al., 2006). A patient safety model (System Engineering

Initiative for Patient Safety) presents these socio-technical categories and their interactions as

inputs to health care delivery processes and recipients of feedback loops from patient and

organizational outcomes (Carayon et al., 2006). Nemeth et al. (2008) offer a conceptual

framework for guiding rapid organizational change once the change levers are determined. The

framework is one tactic for increasing the adaptive reserve of a health care organization (Nutting

et al., 2009). Adaptive reserve is used to describe a clinic's ability to keep pace with rapid

development and change. Davidson and Chismar (1999) use a model for technology-enabled

organizational transformation to emphasize that technology change must be supportive of,

supported by, and coordinated with changes in organizational structure, strategy, roles and management processes.

New process and organizational characteristics are surfacing in the literature that offers a new lens for understanding the enablers and constraints of change in health care organizations. For example, Gittell et al. (2009) identify relational coordination fostered by high-performance work practices as an organizational characteristic for improved organizational performance. One area to explore is whether process measures like relational coordination (Gittell et al., 2009) are more effective as inputs in improvement strategies than traditional reliance on outcome measures only.

Some technology change models include the Technology Acceptance Model (TAM) and the Technology Use Mediation (TUM). TAM is well known in the IT literature for studying perceived IT usefulness and perceived IT ease of use on actual IT use and recently, has been employed in health care (Holden and Karsh, 2010). TUM offers stages of IT implementation including establishment, reinforcement, adjustment and episodic changes (Orlikowski et al., 1995; Davidson and Chiasson, 2005). All of these approaches suggest focusing on both organizational and technical artifacts to reach health care delivery outcomes. More recent literature suggests not only a focus on (technical and organizational) inputs and outcomes but also on the processes or work systems that deliver care (Crosson et al., 2005; Gittell et al., 2009; Carayon et al., 2006). Our study supports the literature by studying the process of primary care relative to standardization.

1.3 Methodology

As part of a broader study (NSF SES-0826842), we conducted a longitudinal case study, collecting data via interviews of HC providers before and after an EHR implementation. Our

research site, Medical Clinic, employs more than 250 physicians, as well as 1,700 additional employees, practicing in 24 locations in the northeast U.S. Medical Clinic is implementing a popular EHR software package used by a number of other medium to large clinics. At the start of this research, Medical Clinic was completing several early phases of its implementation, including clinic management, telephone messaging, lab order entry and pharmacy systems. This study is examining the effects implementing full EHR capability including the use of computers by care providers in exam rooms.

1.3.1 Data Collection

A three-phase data collection scheme was followed at Medical Clinic. The first phase was conducted before go-live, a second phase was conducted approximately six weeks after imple-mentation, and a third phase was conducted about one year after implementation. We chose five internal medicine primary care sites, sampling from large and small, and urban and suburban sites. We selected five internal medicine physicians as interviewees, one at each site, to provide variance across years of medical experience, experience with computers, as well as supporters and non-supporters of the EHR. We added two additional sites and two internal medicine physicians for round three.   For each physician (PCP), we selected their associated practice manager, and one each of the medical professionals supporting them, e.g., a nurse, a medical assistant (MA), and a patient support specialist (PSS), chosen to ensure appropriate sampling (Pope and Mays, 2006).

Interviews were conducted by the university researchers. Initially, we attended each other's interview sessions to ensure interview consistency beyond what following the same protocols provided. After the first few interviews, each interview was conducted by one university faculty member assisted by a graduate student. To ensure confidentiality, Medical

Clinic's employees did not participate as interviewers or observers and all data were stored and analyzed by the university researchers. Each interview was approximately fifty minutes, and was taped (with permission) and transcribed. All data collection was conducted using IRB-approved procedures. Because Medical Clinic used a site-by-site phased implementation, collecting one round of interviews took place over about five months. We have a total of 78 interviews with the same interviewees completing both the before and after go-live implementation interviews. Because participants were promised confidentiality, we have reported only short quotes that cannot be used to identify the speaker.

1.3.2 Data Coding and Analysis

These interview data were analyzed using grounded theory techniques. Specifically, we followed the advice of Strauss and Corbin (1998), first doing open coding of the data to identify general themes, then doing axial coding to identify relationships among the themes, and finally doing selective coding to generate theoretical themes and compare these themes to the literature. While we followed Strauss and Corbin's coding procedures, we remained true to grounded theory tenets that the data dictate which effects to examine, and researchers do not allow prior expectations to bias their perspective (Glaser and Strauss, 1967).

Using codes the data suggested, two of the university researchers coded interview transcripts as they became available, using the NVivo software package to support and store the coding. Coding of earlier interviews helped us refine our questions for later interviews. To ensure coding consistency, we jointly coded the early interviews. As we coded, we reviewed the resulting coding and discussed new codes emerging from the data to further ensure coding consistency, as well as involve all researchers in thinking about the emerging theoretical insights.

The overall theme of standardization initiated by the EHR emerged from the data. For each major category related to the dynamics of standardization, we performed axial coding. Axial coding focuses on understanding the relationships between the themes identified in open coding and the conditions, causes, actions/interactions, and consequences associated with themes related to standardization. For example, this coding identified a relationship between standardized data and standardization of the physician role that resulted in a concept of increased ease of substitutability amongst physicians. Standardized data were more accessible, enabling higher quality of patient care and therefore, higher physician satisfaction with their jobs.

Table 1.1 shows the coding statistics related to the standardization theme. These statistics represent over 4,600 coded passages in our interviews. The second column lists our original open codes, from which the categories emerged (column 1) as we performed axial and selective coding. The fourth column indicates the number of sources or interviews containing the code. The last column represents the number of total passages coded with the open code.

Table 1.1 Coding Statistics for Key Codes

Category	Property	Number of sources with this code	Number of passages with this code
Primary Care Process	Process Improvement	50	151
	Standardization of Process	10	19
	After Visit	19	34
	Before Visit	20	32
	During Visit	9	18
Standardization of Data	Ease of use	47	135
	Navigation	12	13
	Smart Features	63	137
	Standardization of data	27	60
	System Use	55	138
	Work Around	19	26
	Data Quality	42	68
	Information Availability	76	295
	Information Organization	36	121
Standardization of Task	Changes – Task	24	62
	Coordinate Care – task	16	21
	Doing more stuff	23	53
	Standardization of tasks	47	149
	Task is defined	9	13
	Task Variability	1	1
	Task Completeness	18	30
Standardization of Role	Adding Value in Role Identity	6	15
	Changes – Role	52	132
	Role Image	16	28
	Substitutability	44	90
	Shifting work to others	14	22
	Standardization of roles	47	107
Initiators	Individual-enabled standardization	17	19
	IT-enabled standardization	14	24
	Organization-enabled stand.	18	32
	Site-enabled standardization	24	37
Individual Style	Control	10	11
	Non-linear process of thinking	7	10
	Style	12	19
	Computer Expertise	35	46
	Change	6	9
	Communication mode	3	5
	Motivation	1	1
	Organizational skills	1	1
	Training from residency	2	2

Table 1.1 Continued, Coding Statistics for Key Codes

Site Characteristics	Site Characteristics	39	74
	Workflow Variability	41	78
	Physical Characteristics	3	3
	Small clinic	6	10
Interactions	Information Sharing	46	100
	Standardization	4	4
	Communication	57	121
	Interactions – PSS	8	17
	Interactions-Call Center	2	4
	Interactions-MA	12	18
	Interactions-Nurse	11	17
	Physician-MA Interaction	14	19
	Physician-Nurse Interaction	21	26
	Physician-Patient Interaction	21	39
	Physician-Physician Interaction	4	5
	Physician-PSS Interaction	8	11
	Follow-up Visit	8	11
Patient Complexity	Patient Visit-Physician	12	14
	Type of Patient	21	40
	Type of Visit	16	26
	Visit Frequency	2	2
	Process of Patient Care	1	1
	Patient Education	18	28
	Patient Forms	9	12
	Decision Making	12	15
	Non-linear process	4	6
	Physician-focused process	1	1
	Capacity	32	70
Outcomes	Efficiency-Clinics	47	101
	Efficiency-Exam Room	21	48
	Efficiency-Lab work	14	20
	Efficiency-Resource Management	33	56
	Efficiency-Utilization	24	41
	Quality of Care	59	115
	Quality of Chronic Care	25	30
	Satisfaction	3	3
	Satisfaction-MA	28	39
	Satisfaction-Nurse	32	49
	Satisfaction-Patient	65	159
	Satisfaction-Physician	36	98
	Satisfaction-Practice Manager	4	4
	Satisfaction-PSS	28	39
	Trade-offs	6	7
	Time	32	97
	Workload	52	134

During the process of axial and selective coding, many relationships surfaced. Table 1.2 shows the axial coding statistics for key relationships across the dimensions of the open codes surfacing in the data. The first column indicates the categories or properties that are associated across their dimensions. The second column indicates the number of interviews or sources coded with the relationship. The last column indicates the number of total passages or coded areas that were coded with the relationship.

Finally, we conducted selective coding to identify theoretical patterns. Selective coding continues a process of constant comparison among similarly coded passages and relationships and involves a comparison to the literature. This process led to a more theoretical understanding of our substantive codes in terms of developing a model, which is presented in the results below. In addition, through this process of constant comparison of the similarities and differences of passages that we had identified as part of our phenomenon, an understanding of the complex relationships across conditions and consequences emerged.

Table 1.2 Axial Coding Statistics for Relationships

	Number of sources	**Number of passages**
EHR Embedded Structure (Associated) Changes - Role	9	9
EHR Embedded Structure (Associated) Satisfaction	6	11
Computer Expertise (Associated) Satisfaction	8	9
Data Quality (Associated) Style	12	15
Doing more stuff (Associated) Satisfaction	18	22
Efficiency-Clinics (Associated) Data Use	26	39
Efficiency-Clinics (Associated) Information Availability	3	6
Efficiency-Clinics (Associated) Physician-focused process	5	11
Efficiency-Clinics (Associated) Satisfaction	14	21
Efficiency-Clinics (Associated) Standardization	43	88
Loss of Human Interaction (Associated) Standardization	9	10
Non-linear process of thinking (Associated) Style	6	8
Organization-enabled standardization (Associated) Satisfaction	5	7
Physician (Associated) MA	6	6
Physician (Associated) Patient Support Specialists	3	5
Quality of Care (Associated) Information Availability	10	11
Quality of Care (Associated) Standardization	13	17
Satisfaction (Associated) Coordination – process	5	7
Satisfaction (Associated) Individual-enabled standardization	17	22
Satisfaction (Associated) Information Availability	30	53
Satisfaction (Associated) Interaction	4	6
Satisfaction (Associated) Physician-focused process	3	6
Satisfaction (Associated) Quality of Care	6	6
Satisfaction (Associated) Standardization (of tasks)	18	24
Satisfaction-Patient (Associated) Type of Patient	10	10
Site-enabled standardization (Associated) Satisfaction	11	18
Standardization (Associated) Non-linear process	17	26
Standardization (of roles) (Associated) Computer Expertise	7	8
Standardization (of tasks) (Associated) Style	26	41
Style (Associated) Efficiency-Clinics	16	23
Style (Associated) Satisfaction	25	36
Task is defined (Associated) Efficiency-Clinics	16	20
Time (Associated) Standardization (of roles)	10	13
Type of Patient (Associated) Task Variability	22	29

1.4 Standardization Dynamics Model

As a result of our grounded theory approach, an overall model emerged describing influencers, conditions, and outcomes that explain a particular state or extent of standardization in a process. The *process state of standardization* is a characterization of the primary care process in terms of standardization at a particular time. Figure 1.2 represents the dynamics of our emergent model.
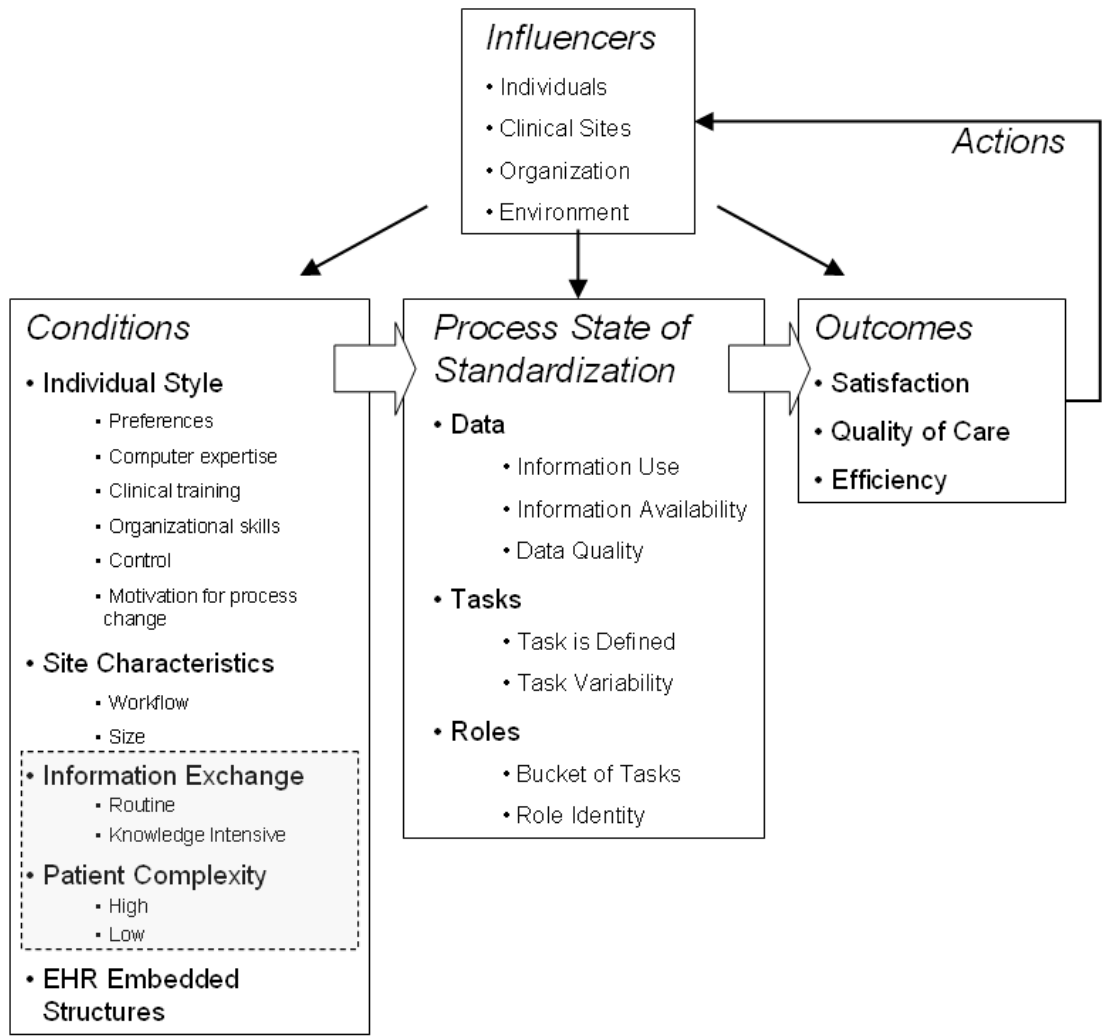
Figure 1.2 Standardization Dynamics Model

In the sections to follow, we describe each component of the model supported by quotations from interviewees. They are quoted verbatim so that the reader can experience the richness of the spoken words.

1.4.1 Influencers

Starting at the top of Figure 1.2, influencers or action takers design strategic or routine tactics that shape how the process state of standardization changes. The strategic actions are purposeful or deliberate acts that are taken to resolve an issue or to reach a goal and in so doing shape the process. Influencers include the individuals working at a primary care site, the clinical site management, the organization or the environment including regulatory bodies and biomedical knowledge as well as the patient and community. This framework and the dynamics between influencers are similar to that suggested by Berwick (2002) for improving health care quality, except that our model suggests an individual level of influence. For example one physician stated,

> I, within a week, was not dictating any notes any longer, which for me was a major goal … I was determined that I was not going, once this was in, that I was not going to dictate. (PCP, site A, round 3)

Other influences might have similar or contradictory goals in driving the process state of standardization. One practice Manager explains the direction set by the organization,

> But they want us to have a semblance of order, so anybody can go work anywhere, and any doc can go to any site, and everybody pretty much does the same thing. (Practice Manager, site G, round 2)

Understanding how the relationship between these influencers affects the process state and outcomes requires taking a deeper look at the intervening and contextual conditions of the environment.

1.4.2 Conditions

Conditions help to explain why different patterns of standardization emerge from the actions of the influencers. Conditions define characteristics along their relationships, which come together to produce a specific situation that explains what is happening with the process state. Conditions are not meant to define cause and effect. They may change over time, affect one another, and combine in various ways across properties (Strauss and Corbin, 1998). The model in Figure 1.2 lists five conditions and their properties that surfaced in the data and influenced the extent of standardization, including individual style, site characteristics, information exchange, patient complexity and the IT embedded structures.

The *individual style* of a physician can shape how clinical tasks are performed, appointments are scheduled, patients are roomed, forms are printed and communication with clinical and nonclinical staff is done. One physician explains,

> Main features of things that are important are pretty similar for all internists, but exactly the way that it's done, the way that you think, the way that you organize, this is very different. (PCP, site J, round 1)

Properties of individual style surfacing in the data include preferences (e.g., communication), computer expertise, clinical training, organizational skills, control, and motivation for process change. The degree of influence of style on the process state depends on the specific role (e.g., physician, nurse, medical assistant, secretary) but is also interrelated with other conditions.

Another condition is the *site characteristics* with properties of size, workflow interrelated

with patient complexity. The differences in site properties in combination with the other

conditions can explain to some degree the emerging patterns of process standardization. One

practice manager explains,

> I think things were pretty standard to begin with before [EHR]. What I find with the
> [EHR] is that you have to manage its functionality and how you're going to make it
> work for your practice. Every practice is different, the needs are different, the
> providers are different, their preferences are different, and their expectations are
> different. (Practice Manager, site F, round 3)

> *Information Exchange* is information (e.g., patient data, diagnosis and treatment plans)

exchanged between roles (e.g., physicians, medical assistants, nurses, secretaries, patients)

during the process of an exam room visit.   The properties of information exchange, routine and

knowledge intensive, are bounded by the state of biomedical knowledge, treatment regimen

selection, standardized medical protocols, complexity of the patient and the state of the process.

Routine information exchange occurs when the data and tasks are predictable to support a patient

plan. One medical assistant discusses a routine information exchange with a physician after the

EHR implementation.

> I like everything. [LAUGHTER] I like everything because there are several doctors
> that you could not make out their handwriting. And now with all the letters and
> everything done through the computer, it's a lot easier that way. I like everything about
> it. (MA, site J, round 3)

> Knowledge intensive information exchange occurs when the data and tasks are not

routine and must follow a knowledge intensive path prescribed by the provider. One practice

manager shares circumstances that require a nurse to use judgment when triaging:

> There are times, like with the nurse on triage on phone who needs to use her judgment,
> and they need to gather information and make the best decision, whether it's a clinical
> decision or do I put this patient in now, or do I wait ten minutes? So they need to have
> the common sense and be able to problem solve and look at the whole picture. (Practice
> manager, site J, round 1)

Information exchanges between the patient, clinical and nonclinical staff occur over the course of the primary care visit representing communication and across visits, specialists and ancillary functions representing coordination of care.

Patient complexity defines a patient in terms of the type of patient (e.g., healthy, chronic conditions), type of visit (e.g., acute, chronic, preventive), patient demographics (e.g., socioeconomic, age) and preferences (e.g., treatment regimens, location of services, education materials, communication). The aggregate level of patient complexity can vary across physicians and sites, contributing to the emerging patterns in the process state. One physician commented,

> I mean part of the danger of predone text is that every patient starts to look the same if you're not careful. (PCP, site G, phase 3)

The *EHR embedded structures* are the data, routines and roles defined in the EHR system as configured by the organization, which are then used by individuals who can then to some degree customize to fit their style. The embedded structures can influence how the tasks are performed, in what order, who enters data, when the data is entered, where the data is entered and how the data is entered. EHRs enable information availability and accuracy by requiring the entry of information at the point of initiation, thereby eliminating the number of information exchanges and shifting the point of initiation toward the provider and away from the patient and staff. One physician notes the complexity,

> And one of the challenges that [the EHR] provides is when you try and put a standardized system on a non-standardized process it creates some friction, to put it mildly. (Physician, site A, round 2)

1.4.3 Process State of Standardization

The process state of standardization is a description of the primary care process at a point in time framed in terms of what is standardized and the extent of standardization. The process state of standardization is a phenomenon that surfaced in the data and is the lens that we employ while studying the primary care process. We define standardization in terms of what is standardized, with subcategories related to data, tasks and roles, and their properties and dimensions. Table 1.3 provides quotes that illustrate each subcategory. A more detailed discussion of each standardization dimension can be found in Appendix A

Standardization of data refers to standardizing the data fields that should be entered, data values that can be entered, and the form and content of data outputs. We organize our data standardization results into the properties of data quality, information use, and information availability.

Standardization of tasks refers to workflows and processes, including what tasks are done, whether the right tasks are done, and whether tasks are done in the same way. The embedded scripts in the EHR software, which dictate the tasks of medical staff and in some cases the sequence of tasks and workflow, strongly influences the extent to which the process is defined and the impact on process variability.

Standardization of roles is associated with increased substitutability between medical staff members and increased use of best practices for the delivery of care. Bucket of tasks and role identity are two dimensions of role standardization that informants discussed.

.

Table 1.3 Standardization Description with Coding Examples

Code Group	Standardization properties	Effects on Outcomes
Standardization of Data	**Data Quality**	**Secretarial Staff (site C, round 2):** "Not a question anymore about what the doctor wants because he has to put it in the system." **PCP (site A, round 2):** "Messages from nurses are better. It has forced them to ask standardized questions; with responses taken down in an order...My instructions back to the nurses are better."
	Information Use	**Practice Manager (site C, round 1):** "some of the DOT phrases I just love. When you're putting in a message and you hit dot-PH, and the person's phone number populates right into your note without having to look it up or write it down, you just verify it with the patient." **PCP (site F, round 3):** "if I'm seeing a diabetic, you click on the diabetes smart set, and it gives you all of their relevant numbers regarding the diabetes. Or if they're anemic, it just, you can bring down all of the relevant lab work to that. So that helps in managing our chronic illnesses. **Practice manager (site J, round 1):** "It's made problem solving a lot easier in my job. It's made not just problem solving like a patient calls and says, this is the problem, or I talked to somebody yesterday, and I don't know where things are at. So for me to be able to jump in from a management point of view and say, OK, well let me see what's going on, I can fix that much easier. I can resolve those issues while the person's on the phone, as opposed to having to pull the chart and review the notes and try to figure out who said this and who signed. You know what? Much, much easier. So yeah, problem solving, identifying what's what. But it's also a watch, I don't mean this in a negative way, it's really helped me identify where staff's growth areas are."
	Information Availability	**Nurse (site C, round 3):** "I think it's wonderful because it's, everything I need is right there. I don't have to leave that exam room. I don't have to go call for charts. I can find what I'm looking for from appointments to when their last EKG was six years ago. So it is such a huge, huge difference." **Practice Manager (site B, round 3):** "the information is available to them, it's at their fingertips, they don't have to run up and down the hallways anymore looking for the doctor to say, what does this say?"
Standardization of Tasks	**Process is Defined**	**PCP (site I, round 1):** "Eventually when I pull up a patient, it'll warn me what needs to be done, the best practices and that will be of great help." **PCP (site A, round 2):** "It has made very explicit what needs to be done, but probably the biggest single change is that when the messages come in, it's very clear on the screen there that you have prescriptions to fill. You have patient calls to answer. You know exactly what you have to do, and whether it's been accomplished. There are no lost pieces of paper." **PCP (site I, round 1):** "I think eventually that would be a reasonable goal that say, the patient prep is more consistent and standardized because they have to go through the [templates] in [the EHR]. But, as you know, it's to get any large group to agree on certain processes is hard, and I think what we need to really focus on is the outcomes, and I think the processes will then fall in place."
	Process Variability	**MA (site F, round 2):** "Every patient is different, with different concerns, different diagnoses." **Nurse (site C, round 2):** "the problem is triaging a patient is not scripted. I'm not going to sit there and ask all of those questions and get those answers. You can get them from a story. You can get them from their conversation, and it doesn't come in the order of, so [templates] are not good for triaging." **PCP (site A, round 1):** "My patients are old, multi-problem, so [templates] don't work". **PCP (site J, round 1):** "Main features of things that are important are pretty similar for all internists, but exactly the way that it's done, the way that you think, the way that you organize, this is very different."

Table 1.3 Continued, Standardization Description with Coding Examples

Standardization of Roles	Bucket of Tasks	**Secretarial Staff (site J, round 2):** "Doctors are booking appointments and labs in their exam room, wow. How many secretaries do you need?" **Practice Manager (site C, round 3):** "But the system states that the medical assistant does the chief complaint, does the social history, puts in any orders for medications, any comment field about any over-the-counter medications that they're on; they never did that before." **Practice Manager (site D, round 3):** "but they want us to have a semblance of order, so anybody can go work anywhere, and any doc can go to any site, and everybody pretty much does the same thing."
	Role Identity	**Nurse (site A, round 2):** "Glued to a chair, staring at a computer, and it is dissatisfying. I love to be on my feet and using my hands and being clinical." **PCP (site C, round 2):** "I feel that we're spending an awful lot of time looking at that computer, setting up dates of return and labs and things, that, and I'm not doing the patient care." **PCP (site H, round 3):** "I did not go to medical school to become a typist."

1.4.4 Outcomes

Using grounded theory, a phenomenon such as standardization should be examined contextually or amongst conditions in which it is embedded, and related to actions through to their consequences or outcomes. Figure 1.2 lists three outcomes that have surfaced in the data and have also been studied carefully in the literature (Bates et al., 1994; Linder et al., 2007; Pizziferri et al., 2005). In this study, we used process examples and impressions from interviewees to get a sense of the improvement in quality of care, efficiency and satisfaction of the primary care process. Examples include the quality of encounter documentation to capture the essence of the patient problem, time in and outside the exam room to complete the EHR transactions, accuracy and timeliness of data for clinical decision making and the accurate and timely information exchanged between tasks and roles.

Satisfaction refers to the degree in which patients, medical staff and physicians are content with the expectations of the delivery of care processes and outcomes. The EHR affords choices to accommodate the variability and complexity of patient preferences, values, and needs, as well as the individual preferences and work styles of care providers.

Quality of care refers to "the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge" (IOM, 2001). For this study, quality of care refers to the perceived quality delivered to the patient during the primary care visit in establishing and communicating the patient care plan, rather than a traditional quantitative measure (such as those defined by Healthcare Effectiveness Data and Information Set (NCQA, 2007)).

Efficiency refers to the utilization of resources and processes to get the best return on investment or dollars spent in delivering high quality care. We view efficiency in terms of process efficiency, those tasks and workflows required to deliver high quality care to patients, as well as resource efficiency, the medical administration and equipment required to deliver high quality care to patients. In this study, efficiency has properties of capacity, consistency, and time. Capacity is defined as the perceived number of patients seen by physicians per day. Consistency refers to the consistency that tasks allow for patient flow during a primary care visit and is related to the quality of the patient care plan. Time is defined as the required time to complete tasks, including personal time, in order to deliver the care needed for the daily patient schedule.

Tension develops when the process state of standardization does not meet the needs of the individual, site or organization causing unintentional consequences in one or a combination of satisfaction, quality of care and efficiency. For example, nurses often used free text for their triaging notes instead of using the recommended template, which in some cases caused tension when clarification or additional information was needed by physicians. Physicians differed in style when assigning tasks to and creating information exchanges with MAs and nurses making substitutability across roles difficult. For example, school forms must be signed by the

physician. Some physicians print the forms in the exam room, sign and hand to the patients.

Others expect the MA to print the form and then find the physician to sign and hand deliver back

to the patient. To address the need for such variations, Medical Clinic has task forces and

forums whose goal is to gain consensus from physicians about standardizing tasks and

information exchanges between roles so as to minimize workflow interruptions and

inconsistencies.

Satisfaction and efficiency were the most immediately observed and easy to measure

unintended consequences after the EHR implementation at Medical Clinic.   Tension occurred

after the EHR implementation, when the process state of standardization changed in terms of the

'bucket of tasks' assigned to roles changed. For example, physicians were required to enter

orders in the system (versus written scripts) and medical assistants were required to enter patient

history, allergies and medications in the system (versus vitals only in the paper chart). The

consequences of the EHR implementation heavily affected physician satisfaction in the time to

deliver care and the perceived level of quality delivered to the patient. Using the model in Figure

1.2 as a lens, the immediate outcome of introducing the EHR embedded structures as a condition

into the process state was the time required by physicians to deliver care to patients. One

physician commented,

> Now every single prescription comes to me independently and although [the nurses]
> have done all of that work, their work hasn't changed. Now it comes to me
> individually, and one could decide well I'm just going to sit here and go ch-ch-ch-ch,
> click 20 and get 20 clicks done in 20 seconds. But the reality is you don't do that,
> you look, because they're all individual now. At the end of that one could argue, well
> it has forced me to look at every prescription, to be absolutely sure, and I can
> understand saying that but in the reality of the situation. Yes, I have looked at all of
> them, but I didn't do any more work than I did before. I did not go into the charts to
> see if their blood sugar or their potassium, or their lipids or whatever were checked
> recently, because the nurse has already done that. And she has indicated that to me
> by sending me the prescription. So I'm actually just taking more time to do the same

amount of work that I did before. So yes, my role has changed. (PCP, site D, round 3)

Physician satisfaction declined as they struggled to return to their previous daily capacity with the same level of quality care. One physician explains,

> I feel that we're spending an awful lot of time looking at that computer, setting up dates of return and labs and things, that, and I'm not doing the patient care. (PCP, site C, round 2)

1.4.5 Dynamics

As shown in Figure 1.2, the process state of standardization is dynamic, changing over time. The process state of standardization evolves, as conditions change through actions taken by the influencers. Such actions are taken in response to outcomes as well as opportunities for improvement. We observed that clinical and nonclinical satisfaction was the direct output of the perceived performance of quality of patient care and efficiency and was the catalyst for process and IT changes at the individual, clinic and organizational levels.

Individuals took action in response to the EHR embedded structures to learn the system and minimize unintended consequences. Problems arose when users could not find the information needed to complete the embedded routines. For instance, one MA explains a workaround necessary for the exchange of accurate information to the physician and to advance the embedded EHR routine:

> Sick visits can be a little bit more complicated because it's hard to find the chief complaint that you're looking for and it can take a little bit longer than just writing down whatever the chief complaint is that someone had told you word for word… I had a ganglion cyst once… I don't want to put cyst because I'm not a doctor and I can't diagnose that it is a cyst. So in the chart I could write, patient thinks it might be whatever. But I can't get that. So then if I just put in something more generalized, then the provider has to go in and change that so he can get the [template] that he needs. It's like that extra step. (MA, site B, round 2)

Over time, individual actions switched from addressing problems to designing for process improvement. For instance, individuals were encouraged to build or request customized

templates to support their work flow and patient population. The availability of data in the form

of audit trails was also used by all staff as a way to check the status or completion of tasks to

support patient care plans and responsiveness to patient needs.

Actions by the organization and clinical sites were initiated to lessen the burden on the

physician. The organization distributed typing software before the EHR implementation and

trainers were available to help them customize their EHR embedded features after the EHR

implementation. The clinical sites also initiated actions as explained by one practice manager:

> Unfortunately, doctors feel like they're now doing the ordering that used to be done
> by PSSs. So that's the negative side to that. But what we've tried to do is put
> processes in place where we're taking other clerical things off their plate that they
> used to do, like opening and sorting their mail, taking care of prescriptions, filling out
> forms more completely before they get them. (Practice Manager, site J, round 3)

As conditions and consequences evolve over time, actions arise in response as shown in

Figure 1.2. Although we have presented these constructs as following a linear path, they rarely

do. This is too simplistic an explanation of events and is not consistent with real life (Strauss

and Corbin, 1998). Instead, actions taken by influencers might be in response to multiple

conditions and consequences. It is even possible that different influencers are addressing the

same consequences but in different and some times contradictory ways.

The above examples provide a small sample of the dynamics observed over time at our

field site. Additional examples of the dynamics that occur as influencers take actions in response

to the existing state of standardization and resulting outcomes, as shown in Table 1.4. The types

of actions, which are categorized as IT design or process design, are noted to show the types of

levers used by the influencers to drive change. Following the model in Figure 1.2, influencers

may be at the organization, clinic, or individual level. These examples provide a sense of the

relationships between the conditions or set of circumstances in which the process state of

standardization is embedded and the actions made by influencers to address events that arise

under those conditions.

Table 1.4 Examples of Dynamics across Influencers and Time

Influencer	Conditions/Process State	Quotes	Outcomes/Actions
Before EHR Implementation			
Organization-Process Design	Primary Care Nursing to IT-enabled Pools. More work for PCP to study details of each patient due to canned questions from nurse.	"I certainly feel like the absence of a primary nurse is a major obstacle to productivity" PCP, Site A, R1	Doctor is not satisfied and routes his requests to his primary nurse.
Organization-IT Design	Standardized protocols for Nurse Triage	"I know triaging is nursing, but it's a lot of computer and you do feel like, gee, I don't need to be a nurse to do this. Someone else could probably do this job just as well as me. Especially with the Standard Text. All the questions are there. So, it kind of makes us feel a little threatened that maybe we're not needed that much anymore." Nurse, Site A, R1	Nurses are not satisfied, free text to accommodate patient complexity.
Clinical Site-Process Design	Variability in message routing from MA to PCP influenced by PCP style	"Everyone is different. This doctor wants you to route everything to them, this doctor doesn't want you to route anything to them, this doctor wants to, if we could all be on the same page I think it would work a lot smoother." MA, Site J, R1	MA is trained on PCP style, process inefficient, MA not satisfied
Individuals- IT Design	When looking at the list of encounters, telephone messages and office visits are mixed. List becomes overwhelming to manage.	"So, I've put a filter in that the telephone calls are in a different thing so you can just go to whatever filter you want and all the office visits are right there." Nurse, Site F, R1	Nurse learns filters for efficiency

Table 1.4 continued: Examples of Dynamics across Influencers and Time

Influencer	Conditions/Process State	Quotes	Outcomes/Actions
One Month After EHR Implementation			
Organization-Process Design	Secretaries can no longer enter clinical information into the EHR	"Secretaries couldn't enter anything into the EHR with the prescriptions. Because they didn't want, I guess for an error reason or something like that. " Nurse, Site F, R2	Standard task shifts to MA and nurse
Organization-IT Design	Red dot on the EHR schedule to inform nurse and MAs when the doctor needs assistance in the exam room.	"It's hard to find. If it stayed on the time, we would see the red dots, but now we've got to scroll up and down to try and find the red dots. Now, if they fix that piece, that red dot would be really nice. I thought, you know, you go in and you find out what the doctor wants, and then you're able to go and get it…That's not working." MA, Site J, R2	Doctor has to leave exam room to tell MA and nurses what he/she wants. Adds frustration.
Clinical Site-Process Design	New paper form process and new electronic system creates different work for MAs	"One day a week, we have one MA that does all the paper, so we all have a turn to do the paper, which evens things out." MA, Site J, R2	MAs split up the tasks at this site.
Individuals- IT Design	PCP did not like system built bone density letter.	"I didn't like the letter, the bone density letter, so yes, we did do my own bone density letter." PCP, Site J, R2	PCP built custom letter and is satisfied.
One Year After EHR Implementation			
Organization-Process Design	MA looks up history when rooming patient.	"It's a brand-new process. We didn't before have to look at previous Pap smears or, you know, previous biopsies or the history, and now we do." PSS, Site A, R3	MA satisfied with helping PCPs.
Organization-IT Design	Short cuts introduced to staff by IT employees	"But now there's a text they taught us recently, dot form and that comes up and we can just F2 everything and put it in. Like, it's progressing. Each month something else comes out and we are able to do something more. It's getting easier I should say." PSS, site I, R3	PSS satisfied with IT enablers.
Clinical Site-IT Design	Nurses use templates to assist PCP	"But they (PCPs) all have their preference list. Their preferences are in there. So you punch in their numbers, and their preferences come up for UTI, for a cold, for upper respiratory. It all comes up, and then you can pend the medication for the physician to make it easy for them. So that if they choose to do that, it's already there. We kind of run our own UTI clinic. You know, a patient calls in. We have the urines done downstairs. We look them up. We pend the medication. The doctor signs it, and off it goes. So it's made life a lot easier for the physicians." Nurse, site I, R3	Process design adds efficiency and satisfaction for nurses and PCP.
Individuals- IT Design	PCP goal to not use dictation services	"I made a conscious decision, first of all, I'm a good dictator but I am a late dictator, so I wanted to get rid of dictation." PCP, site A, R3	PCP creates short text for efficiency

1.4.6 Model Discussion

Through axial and selective coding, we were able to trace nonlinear connections among conditions, actions and outcomes that have shaped the standardization of the process state. We present a model that offers a view of the dynamics shaping the process state of standardization at the microsystem level, encompassing all levels of influencers. This model is unique in that it combines models of the individual journey (Karsh, 2004; Nemeth et al., 2008; Kaplan, 2001) with system models (Berwick, 2002; Carayon et al., 2006; Wagner, 1998) to support organizational decision making regarding IT and process design at the individual level. High level models are useful in understanding how the pieces of a system (e.g., organizational systems, IT, patient and community) may interact to produce better outcomes, but do not necessarily provide insight as to how those outcomes can be achieved through actions at the microsystem and individual level. Additionally, while the body of knowledge for IT implementation (Holden and Karsh, 2010; Orlikowski et al., 1995; Davidson and Chiasson) and process redesign (Parks, 2003; Sharrock, 2007; Wood, 2004) is extensive, and each recognize the need for the other, very few studies connect the two organizational change enablers into one microsystem level model (Gittell et al., 2009; Carayon et al., 2006). Our model is unique in that it employs a standardization lens at the microsystem level.

Acknowledging the individual journey in shaping the process state of standardization is important because it is at the individual level that actions are taken that collectively can enable the realization of organizational goals. The model highlights that each individual journey is unique because conditions for each individual are different. That is, the individual style and the clinic site characteristics can shape to what extent standardization is adopted by the individual user. Organizations recognizing the individual journey can provide solutions to eliminate unintentional consequences and align individual actions to support organizational goals.

The model also suggests that there are other conditions that shape the extent in which embedded EHR structures and process designs are used to standardize a primary care process. Patient complexity and information exchange independently and collectively dictate the extent to which standardization can be actualized. Tension arose after implementation at Medical Clinic when the process state of standardization did not support the needs of complex patients and the exchange of information to support those needs. For example, nurses and physicians were asked to use templates for documentation of patient encounters. The templates as implemented supported simple acute visits or physicals, but required a lot of effort and time to fit to multi-symptom or multi-chronic disease patients. Physicians and nurses reacted with workarounds, using different forms of free texting or dictation. In doing so, consistency across roles to standardize those tasks was diminished. That is, the ability to reach a desired process state of standardization may not be realized within a site or organization if users within the same role choose to complete tasks associated with different degrees of standardization.

The model contributes to a better knowledge of the set of circumstances or conditions influencing the process state of standardization and the consequences and actions that arise when the process state is misaligned with the needs and goals of the influencers. Having this insight, managers can make better decisions before the implementation of IT and process redesign efforts to minimize unintentional consequences by evaluating salient conditions and moving users as quickly as possible along their individual journey to help realize organizational goals.

1.5 Pathways for Balancing Standardization and Flexibility

A pathway is a sequence of information exchanges across tasks and roles to achieve an end result or to meet the needs of a patient care plan. A pathway follows a physician's process of thinking to satisfy a care plan for a patient. For an exam room visit, one pathway can be a

predefined list of tasks and information exchanges between the physician and staff to complete a well visit for a healthy 20-year old patient. However, even with a predefined visit type, a pathway is typically not established until the physician determines a diagnosis or reason for visit based on information exchanges and a care plan based on patient preferences. The level of patient complexity is somewhat known prior to the exam room visit based on the medical history of the patient and the type of visit scheduled. However, it is at the point of information exchange between the MA, physician and patient that the future pathway is intuitively identified by the physician and next steps are initiated.

Pathways can be defined in many ways. For example, the delivery process might be designed or thought about as if there is one pathway – an exam visit, or by a disease (flow chart) or chief complaint. There are many pathways, typically defined clinically. The view that we take here is from a process design and standardization perspective. This is similar in some ways to determining value streams in lean (Womack and Jones, 1996), but different in that we are not dividing resources in primary care around these paths. The final process design for visits must accommodate all pathways, because the path is not known until the physician determines it.

1.5.1 Pathways

We uncovered a matrix of conditions that shaped four distinct pathways related to balancing standardization and flexibility, as reflected in the process state of standardization. Figure 1.3 shows the conditions of information exchange and patient complexity intersecting to create four pathways. The conditions of physician style, site characteristics and the EHR embedded structures from Figure 1.2 also intersect and shape each pathway, and the resulting process state of standardization, and represent levers for the influencers to consider when designing the pathways.

Figure 1.3 Matrix of Conditions Shaping Four Value Steams of the Process State

We focus on those information exchanges involving primary care physicians near the identification step of the pathway where our data is the richest. However, information exchanges between each clinical and non-clinical role can be discussed within this matrix. Examples of the pathway conditions, consequences and influencer actions across various information exchanges are provided in Appendix B.

1.5.1.1 Pathway of Standardization

The standardization pathway occurs when the patient complexity is low and routine information exchanges exist and are a candidate for process standardization. This pathway can be designed to support standardized care plans and the ability to embed decision trees for patient specific needs and preferences. EHR embedded elements when designed to meet the needs of care delivery can enable a high level of standardization across data, tasks and roles and hence, substitutability and consistency across sites, leading to efficiency, quality of care and satisfaction. At Medical Clinic, physicians were encouraged to use the system templates for routine chief complaints and were discouraged to use their own style (e.g., dictation) in performing these tasks.   One physician shares his view on quality of care.

> We feel it's a better outcome with [the EHR] because when we do the [templates], everything is there. So the labs, pertaining labs, pertaining testing, everything is there. So you don't miss things. It's all in one order. So actually that also helps with the patient care. It's reminding the physicians, OK, you didn't do this. So at least

you can think about, OK, I need to order this. I think that's another good feature.
(PCP, site H, round 3)

While standardization of data and tasks has its appeal, how the data is structured by the system is not always intuitive to the users. Interviewees criticized the embedded data as either too rigorous or offering too many choices thereby affecting efficiency. Extra time is spent either scrolling through too many options or entering additional search criteria to find the desired data. Issues in this pathway typically arose immediately after EHR cutover and were quickly rectified through actions by the influencers to modify the IT system. For instance, one physician talks about his dissatisfaction with the time to find data immediately after cutover.

> The first time I ordered, tried to order an Aerochamber, a list of about 25 options came up. Now, I honestly didn't know there were that many Aerochambers in the world. And I couldn't pick out the one I wanted. (PCP, site G, round 2)

The upfront design of the IT system can minimize this type of unintended consequence by involving medical staff in the verification of lists as done at Medical Clinic. Actions by the organization and clinical sites also shaped the information exchanges between physician and staff, further enabling the extent of standardization in this pathway. For instance the medical assistants were asked to perform more clinical duties like a review of historical procedures. One MA commented,

> Lately, we've been doing more background. It's a brand-new process. We didn't before have to look at previous Pap smears or, you know, previous biopsies or the history, and now we do. It makes for a more accurate current Pap smear, but it wasn't implemented before. It's brand-new. (MA, site A, round 3)

Over time, the process state of standardization of this pathway changed to reflect a higher level of standardization as the conditions of style and embedded structures changed and as different individuals became more familiar with the system. For instance, nurses at one site initiated new ways to offload the physician by streamlining information for routine complaints. One nurse noted,

No longer do they have to see a UTI patient. We take care of that issue. All they have to do is sign it, and send it once we get the information. They no longer have to look anything up. If I'm already looking the lab up, why don't I copy, and paste it, and put it on my telephone message? And then they can see everything right in the message. So we have made it a lot easier for them. (Nurse, site F, round 3)

In return, information exchanges between the nurses and physicians were more structured.

One physician commented on the quality and efficiency of such exchanges.

Messages from nurses are better. It has forced them to ask standardized questions; with responses taken down in an order...My instructions back to the nurses are better. (PCP, site A, round 2)

While this pathway is easy to implement and modify and is associated with high outcome

levels, it covers fewer than 50% of all primary care visits (Green et al., 2006). Patients having

knowledge intensive problems or having one or more chronic conditions require some level of

flexibility to document and communicate the patient diagnosis and care plan. And interventions

like EHRs are most effective in structuring flows that occur with low to moderate patient

complexity and routine information exchanges. Tension and dissatisfaction arise when a one-

size-fits-all approach with rigid EHR embedded structures is implemented for all four pathways.

1.5.1.2 Pathway of Multiple Standardized Paths

The Multiple Standardized Paths pathway occurs when conditions of high patient

complexity and routine information exchanges exist and is a candidate for standardization based

on flow sheets and other approaches for managing chronic care conditions. This pathway is

categorized by the use of multiple standardization tools during routine visits (e.g., acute care,

annual physicals, follow-up on chronic state).   A simple acute care visit for a relatively healthy

patient can become a complex decision tree for a chronically ill patient. These standardization

tools were initially overwhelming for medical staff to use resulting in an unintended

consequence of dictating. One-on-one training immediately after the EHR cutover increased use

of user-defined shorthand methods, counteracting issues in efficiency. For instance, when a

patient with diabetes has a preventive or acute care visit, the type of visit might be routine but the

creation of the patient's care plan would require additional steps for verifying the current state of

the chronic condition. One physician commented on an IT enabler,

> I think one thing that looks like it will be really helpful is they have these flow charts
> so if you have a diabetic patient, you click on the diabetes flow chart and
> automatically it brings up all their blood sugars, their hemoglobin, A1C's, their
> cholesterols. (PCP, site F, round 1)

Individual physicians have also influenced the pathway by implementing local templates to

streamline more complex patients through the care delivery process. One physician noted,

> And I see a lot of chronic disease, so I have actually learned how to bend the system
> to that type of patient. So for example, if you come to see me and it's just a simple
> high blood pressure, you have high blood pressure, see you once or twice a year
> whatever. I now actually have my notes that I have formatted for that visit, so I pull
> out my note, pops in there, and there's a few questions that I have to answer, a few
> things I have to change. So I've bent the system to that, so that makes things very,
> very easy. (PCP, site D, round 3)

The organization as an influencer has also enabled EHR protocols as a reminder and

standardization tool for clinical staff to perform tasks to aid the physician on routine tasks for

more complex patients. One practice manager commented,

> As the MA puts in, there's certain things that they have to, you know do to make sure
> that it's either the pulse ox or the inhaler to make sure that their peak flow is good.
> All those same things are standardized, so that has been very good because we used to
> have that all over the board. Some people would do it. Some people wouldn't do it.
> Some people required it. Some people didn't require it. So that has definitely, you
> know, those questions come up, have you done this? Have you done that? So that is
> good. So that piece, I think, is better. (Practice Manager, site E, round 3)

Over time this pathway took shape from a more flexible state to one of many

standardized paths. While this pathway can continue to become more standardized by adding

additional standardized paths for more chronic conditions, the number of patients with multiple

chronic diseases and related comorbid conditions continues to grow along with biomedical

information to treat and monitor their conditions. Additionally, most care delivered to chronically ill patients is administered by the patients and their families requiring a degree of collaboration in the management of outcomes. Patients also vary greatly in motivation, the amount of education they want and the level of control in decision making. This condition requires flexibility for the physician in assessing their current state and designing care plans for their preferences.

1.5.1.3 Pathway of Structure with Flexibility

The Structure with Flexibility pathway occurs when conditions of low patient complexity and knowledge intensive information exchanges exist and are a candidate for structure with a physician focus or flexibility. This pathway is categorized by modifying or augmenting an existing more general template(s) for the unique chief complaint(s) or dictating the patient story. For instance, when a relatively healthy patient has a visit for a rare short term illness or nonstandard complaint, the physician might have to conduct some medical research or bend a visit structure to meet the unique needs of the patient. In this case, there are no predetermined care templates to use. To support physician satisfaction and efficiency, physicians were allowed to select their level of standardization, e.g., they could continue to use dictation services instead of using visit templates. One physician explains the challenge in capturing the essence of the patient complaints in a system full of isolated templates.

> They don't tell you that in the beginning. I'm here for the back pain, so you put the back pain, you open the Smart Set for back pain, and you do the back pain thing. And then say, and what about my rash? And remember my heartburn, and you know my migraines are really not getting that better anymore. So now you really should have put in the back pain and the headaches, and the rash, and UTI, or whatever, so then you have like a 20 Smart set and although you can open only one Smart set at a time. So then you click to the back pain Smart set, you got it out then go back, and then open the rash Smart set and then you order the skin test and whatever. So it takes you longer. And I don't like the progress note still, because the progress note

for the back pain will take me a much longer time because they have to have all the categories of back pain. It's not this was rather simple, she was shoveling the snow, there's nothing on exam, but still to document that on the EHR note will take me much longer, so I'm better off dictating, so I'm still dictating. (PCP, site J, round 3)

The organization has provided some level of assistance in capturing the patient data and the information exchanges between the physician and the MA. This is one mechanism to flag the physician of hidden causes. One MA commented,

It's if you put in a blood pressure that's over 140 over 80; it stops you and asks you is this correct? And it's very rare that somebody, you know, we have a lot of patients with hypertension so, you know, it's just that redundancy. But it is a precaution and it is making you double-check yourself. So it's looking out for human error so I guess it's, it does it with the weight also. If you put in a weight that's over a certain amount, a percentage of their last visit, it alerts you, which, again, it's looking for human error, you know, which can happen. So it's kind of a love/hate kind of thing because it does slow you down and you need to get out, but you need to be accurate at the same time. (MA, site A, round 3)

Physicians complain that the output of the system becomes difficult to understand when covering other physicians' patients or even reviewing their own notes. One physician discussed,

I'm not doing the patient care. Then going out and trying to use the system the Smart Sets which, you're using three or four Smart Sets if you get somebody in with different things. And then when you try to read them afterwards, they don't make sense. There's not a flow, there's not a, I could read the note, and I understood what I saw the patient for. It's kind of scattered, and there's almost too much that you have to go through. I still don't feel comfortable with it, and so I'm not a fan. (PCP, site C, round 2)

This pathway remains an area where the rigidity of the EHR embedded structures has not allowed for the flexibility of unique individual cases. Each physician has tailored their style to the circumstance making substitutability across roles difficult and communication and coordination of the patient plan less than optimal. In order to capture the knowledge intensive information in this pathway, dictation within the EHR and bending the system to meet the preferences and style of the physician remain as the two main levers in molding this pathway to fit the structure of the EHR embedded structures.

1.5.1.4 Pathway of Flexibility

The Flexibility pathway occurs when conditions of high patient complexity and knowledge intensive information exchanges exist and are a candidate for a highly physician focused or flexible process. This pathway is categorized by allowing physicians to choose the method that they believe best captures the essence of the patient complaint and care plan. Patients in this category typically have more than one chronic disease and related comorbid conditions and are being seen by the primary care physician for related complications with an acute ailment. This pathway clashed radically with the IT embedded structures at the EHR cutover. Multi-symptom patients with several chronic diseases did not fit any single template available in the EHR. Standardized templates created opportunities for error in follow-up and treatment within Medical Clinic and especially when information exchanges occur outside of the organization or EHR system (e.g., specialists, referrals, specialized testing, detailed patient care plans). Short term, some physicians used a single diagnosis template for the major complaint and then used free text the rest of the visit note but most physicians simply dictated the entire encounter. This situation led to inefficiencies and dissatisfaction resulting in some short term resistance and circumvention in utilizing the system. Long term, however, physicians were encouraged by the organization to customize templates for handling common scenarios of complexity within their practice. Eventually, the organization has succumbed to the style, demands and preferences of the physicians recognizing the need for capacity and effectiveness of delivering care. This scenario has left the organization with many variations across physicians in documenting and communicating patient diagnosis and treatment plans.

One physician frames the challenge of using embedded structures for this pathway.

People are very inconsistent about how they do the notes. If you do a canned note from this system and just fulfill your obligations, you'll never get a flavor for what

actually happened to the patient within the interview and the physical exam. If you dictate then you're kind of defeating the purpose of the system, but that remains the best way to get a note that actually reflects nuances of how the patient looked, what the physician who was caring for them was thinking, why they chose to do what they did. So that piece is not uniformly available. If I want to know how your CT scan came out, this is beautiful. If I want to know if you had a colonoscopy and, if so, when and what did they find, perfectly fine, last mammogram, what was on it, last CT scan of the chest, all of that stuff absolutely fine, but not the process of op. But what I like least about it is that there's still too much that's canned and there's not enough support for people who are trying to make it not canned. (PCP, site A, round 3)

Dictation is the most common form of documentation. Recognizing that dictation was necessary to deliver care in this pathway, the organization offered real-time transcription speech recognition software to minimize the transcription time and costs. One physician talks about this opportunity,

Each patient is unique. Especially with a complicated, difficult patient, it's individual. So it's hard to do common notes for those patients. Maybe if we could implement [audio to text] software with this, so we don't have to dictate and then wait for a dictation for a few days. I think that would be a good option. (PCP, site H, round 3)

A MA scribe was a piloted approach at one clinical site to help physicians focus on the complexity of the patient rather than the computer and in particular for those physicians who needed assistance with computer skills. The costs of this program were deemed too high to migrate to other practices. One MA described the experience,

They [MAs] can be the one focusing on the computer while the doctor's still focusing on the patient. Because like I said, if you miss one little category it won't let you continue and then you're looking back to try to figure out what's wrong, why isn't it letting me move on? And that puts the doctor behind, a simple little thing like that, because then they start getting frustrated because the patient's still going, and they're trying to concentrate on them but they're also trying to concentrate well why isn't my computer working for me? So that's where I come in and he doesn't have to worry about all of that. (MA, site D, round 3)

With the growing complexity of delivering care to the chronically ill, this pathway might never evolve to standardizing what is captured during a primary care visit but there might be an opportunity to standardize how the information is captured.

1.5.2 Designing for Pathways

Hall and Johnson (2009) suggest standardizing the process steps before and after a judgment-based worker, an artist or physician, as a mechanism to gain efficiency. We found that it is at the point of information exchanged between clinical and nonclinical roles through tasks and data that make standardization downstream from the physician difficult. The patient complexity (type of patient, type of visit and preferences) and the information exchange (routine or knowledge intensive) determined how much a process could be effectively standardized in terms of data, tasks, roles and formed four distinct pathways (Figure 1.3). The value of such pathways is that they frame choices about how to use embedded EHR structures to support effective patient care plans, going beyond generic advice to redesign processes in conjunction with IT implementation. The delivery process as a whole must be designed to accommodate four pathways, recognizing each pathway can lead to more robust process designs.

To accommodate the ***Standardization*** pathway, the delivery process should incorporate IT and process design choices to promote standardization when conditions of patients with low complexity and routine information exchanges exist. In this pathway, the embedded EHR structure supports standardization, and can help guide the way work processes are designed. Influencers typically take action to mold the conditions to support the structures of the EHR. At Medical Clinic, the organization provided additional training for individuals to use templates and to standardize workflow across sites. One Practice Manager shared a focus on process design to support the information exchange between the PCP and the MA:

> Because the nice thing is that the MA is going to be able to go into the exam room, put in the focus notes, put in the vitals of that patient, secure the computer and when the physician walks in, by putting in their password on top of the MA's, it changes their security level but yet opens up to exactly where the MA is. So a medical assistant can't write a prescription but a doctor can so they'll have all the functionality and security levels that they'll need when they get into that exam room. So I think those process changes will be the single biggest change for managers and determining

how to allocate resources, be they staffing resources or physical resources to best support the practice. (Practice Manager, site C, R1)

Unintended consequences after the EHR implementation typically arose when the EHR embedded structures did not align well with the standardization needs of this pathway. For instance, the seamless flow of information exchanged between the MA and PCP is critical for the PCP to properly diagnose and treat the patient complaints. One MA commented on the incomplete standardized list of chief complaints in the system causing inefficiencies and inaccuracies with information exchanged with the PCPs:

> We don't have the same options that the doctors do as a diagnosis. So if somebody comes in with dysuria, for instance, there's no dysuria for us, even though that is, could very well be what the diagnosis for the visit is going to be. We still don't have that option. So we were typing in UTI because that would come up, but then we're kind of making an assumption that it is a UTI. And a lot of times it's not. So we had one of the doctors come out to us one day and say hey, you know, please don't put UTI in there anymore. And we had to say sorry, that's kind of our only option, you know. (MA, site B, R3)

The pathway requires the EHR embedded structures to be consistent with medical terminology, tasks and roles across routine information exchanges. One PCP describes the unintentional consequences of a mismatch:

> My sense is the program was built by non-medical people, I'm sure it was built by computer programmers, that had some understanding of what we did, but they don't know what we do. So frequently they have put material, they put things in the program where intuitively from my point of view it doesn't belong. Or for example, there's medical terminology, they have a lot of it in, but not a lot. So frequently for example if I see somebody who I think has an enlarged liver, which we call hepatomegaly, if I try to enter hepatomegaly the computer has no idea what I'm talking about. But if I put down large liver, there it is. So small items, but we're trained to think with certain language and terminology, and sometimes you're forced out of that, and sometimes it takes 30 to 45 seconds to realize I better not use my medical terms, I just should go back to plain everyday English, to find what I'm looking for to enter a term or a problem that I'm looking for. So that's a little frustrating for us, and doesn't sound like a lot but it is. (PCP, site D, Round 3)

Missed opportunities to realize the potential standardization level of this pathway occur when routine information exchanges or the ability to capture low patient complexity complaints

become problematic. Continuing to add and improve protocols for clinically known routine acute, preventive and chronic searchable on medical terms would eliminate workarounds and frustrations in this pathway.

The design of the *Multiple Standardized Paths* pathway should involve making IT and process design choices to standardize along paths of chronic care conditions. In this pathway, the EHR embedded structure can be used successfully to guide the process redesign. For instance at Medical Clinic, the organization provided one-on-one training for physicians to use flow charts and customize templates to fit their style, MAs completed chronic care related standardized tasks in rooming the patient and physicians learned to bend the system to fit their needs.

Unintended consequences after the EHR implementation typically arose when the EHR embedded structures did not align well with the flow of multiple chronic care needs of this pathway. One physician describes the rigid structure of the system preventing flow in the exam room:

> Q: How would you have designed it? Just kind of high level, are there things that –
>
> A: There are lots of things that I do over and over again. For example, this is a great example. The vast majority of people that I see in my practice I can categorize of having five or six chronic diseases. And if I listed those I'd probably include 78, 80% of what I do. And when I see those problems, there are certain things that I always do. So somebody has diabetes, when I see them, I always want to know what their blood sugar control has been, that's rote, this is what you do. The system the way it's set up, every time I see that diagnosed, I have to go in and pick the things that I want, as opposed to [my preferences], that this is what I want, and just one click, bang, it's done, for reasons that are beyond my ability to understand, we can't do that, or we won't do that. I don't know if it's we can't, or we won't. (PCP, site D, R3)

Missed opportunities to satisfy the needs of this pathway occur when routine information exchanges to capture high complexity patients become problematic. Developing protocols for evidence-based chronic care illnesses and the ability to connect them in a seamless flow of

information exchanges represent the opportunity to support this pathway in upfront IT and process design.

Designing the *Structure with Flexibility* pathway requires a process design that is not as structured by the IT and allows for the governance of individual style and site characteristics. At Medical Clinic, because patient complexity is low for this pathway, general templates were often used during the exam with documentation through dictation. A practice manager describes the needs of this pathway:

> If there's something that you want to say but it's not exactly like the [template] and you can't find the right word to use [LAUGHTER], so they end up free texting. So I think that's another thing that again, I think it's a whole new way of thinking. (Practice Manager, site A, R2)

Missed opportunities to satisfy the needs of this pathway occur when knowledge intensive information exchanges for relatively healthy patients become problematic. Inefficiencies and frustrations occur when time is spent searching for the right template to capture a unique complaint. One practice manager explains the inefficiency of trying to use a standardized template for an unpredictable patient complaint:

> What used to take one and a half to three minutes is now three to five, so almost double in some cases. It depends on the patient and the reason why they're here. Is the problem easy to pull up in the system? Is it listed under chief complaint when they can get the [template], is it not? (Practice Manager, site J, R2)

Without such a defined process, physicians developed their own methodology through using system generated templates, customized templates or dictation causing inconsistencies across information exchanges. Inconsistency caused frustration and extra time for physicians (and staff) to read and understand knowledge intensive information from other physicians or questions from nurse triage. One nursed shared an improvement effort for information exchanged between nurses and physicians:

They've changed the template to basically put what's your question or disposition at the top. Before it was reversed and it was all the way down the bottom, and it drove physicians crazy, cause they don't really want to read, you know, a million yes-no, yes-no things. Like if it's a cold or a cough or something easy like that, you're not bringing the patient in or, OK, everybody's throwing up now, gastritis, everybody's sick. So if we're doing the basic thing of what to do for that, we're using [templates] and it's really easy. You're telling the patient what to do and then that just is done, and that message you just LOS and follow it up to them. They'll just click it out of their box. The tough thing is if you have the complicated person that has, you know, specific questions, it's impossible to [template] some of those people, and you can't. And I think the clinic realizes that. You have to ask the specific questions of the doc, you know, if you're not going to bring the patient in. (Nurse, site F, R3)

When scenarios arise that do not fit the structures in the EHR, workarounds are initiated to support patient care. For example, MAs might write notes on a paper encounter form, free text in the notes section in the EHR, send an electronic message or verbally communicate to the physician. The lack of a standard workaround causes variation and inconsistency in information exchanges affecting efficiency, satisfaction and quality of care. A flexible and standard methodology to capture unique occurrences is an opportunity to support this pathway in upfront IT and process design.

Effectively managing the *Flexibility* pathway requires a heavy focus on process design with minimal structure by the IT. Some opportunities for standardization exist up or downstream from the physician tasks and the possible use of customized templates. At Medical Clinic, much of the tension after the EHR implementation existed in this pathway as physicians tried to support the overall direction of the organization to use system-built templates or to create customized templates. One physician explained the frustration of fitting templates to multiple symptom patients:

The other thing is that when you populate diagnoses, I've had some people who have 36 diagnoses listed, which one are we here today for? And you go back and put all 36 in, or do you put, so it is very, very frustrating on that part. (PCP, site C, R2)

Unintended consequences resulted in the inability to support all patients through the use of

templates. One practice manager explained the challenge of using templates in this pathway and

the suboptimal results of capturing the essence of the patient visit:

> The doctors have created a lot of their own smart phrases. Again, I think a lot of them
> think it is canned text, and again you can make a smart phrase, and it might work for
> this patient but doesn't quite work for that one, so that's a difficult concept for them
> too. (Practice Manager, site A, R2)

Missed opportunities to satisfy the needs of this pathway occur when knowledge intensive

information exchanges for patients with high complexity become problematic. Medical Clinic

added functionality to standardize how the documentation (e.g., dictation into the system) was

complete supporting the flexibility needs of this pathway. Additional process design initiatives

were implemented at Medical Clinic supporting high complexity patients including streamlining

the information exchanges between the MA and physician. A practice manager summarized the

progression of process designs:

> Well going back in the beginning when the EHR first started there was a lot of, when
> the doctors felt they were doing a lot of secretarial work, and so that was revamped
> and taken back away from the doctors and given to the whoever, MAs, nurses and
> PSSs. And one of the things can help them is the MAs when they go in, and then
> they bring up the medications they're on, and they review it with them, and if the
> patient's no longer on that med they put a check mark, they don't delete it, but they
> put a check mark, so when the doctors goes in and pulls up the meds and said, oh, so
> and so put in the rooms? And the MA said that you're no longer on this and verify,
> and then the [PCP] deletes the meds and the duplicates, but it cleans it up. (Practice
> Manager, site F, R3)

Recognizing and supporting this pathway upfront in IT and process design decisions would

eliminate a lot of frustration, time and ineffectiveness to communicate and document patient care

plans. Even with pathways requiring flexibility, it is still necessary to make process and IT

decisions as to what the structure of the information exchanges and the documentation of such

exchanges should look like.

Most interventions like EHRs focus on the low to moderate patient complexity and routine information exchanges. Tension and dissatisfaction arises when an EHR system with embedded structures for a one-size-fits-all approach is implemented for all four pathways. The contribution of this paper is identifying workflow into four pathways and planning and designing for the individual needs of each pathway upfront.

1.5.2.1 Process Design Tool to Describe Standardization State

We created a process design tool to describe the process state of standardization in terms of data, tasks and roles for a specific site, organization or pathway. A process state of standardization measured after the EHR implementation at Medical Clinic is depicted in Figure 1.4. The diagram illustrates relationships and relative measurements for one clinical site, as developed from our after implementation interview data. The tool is a qualitative relational diagram representing aggregate data and tasks per role. Each clinical site has a different profile based on individual journeys; it might be useful for an organization to study the variation across sites. Full standardization, which is at the center point of the graph, is impractical and undesirable due to innate variability and complexity in patient types, daily workflow interruptions, and patient and staff preferences.
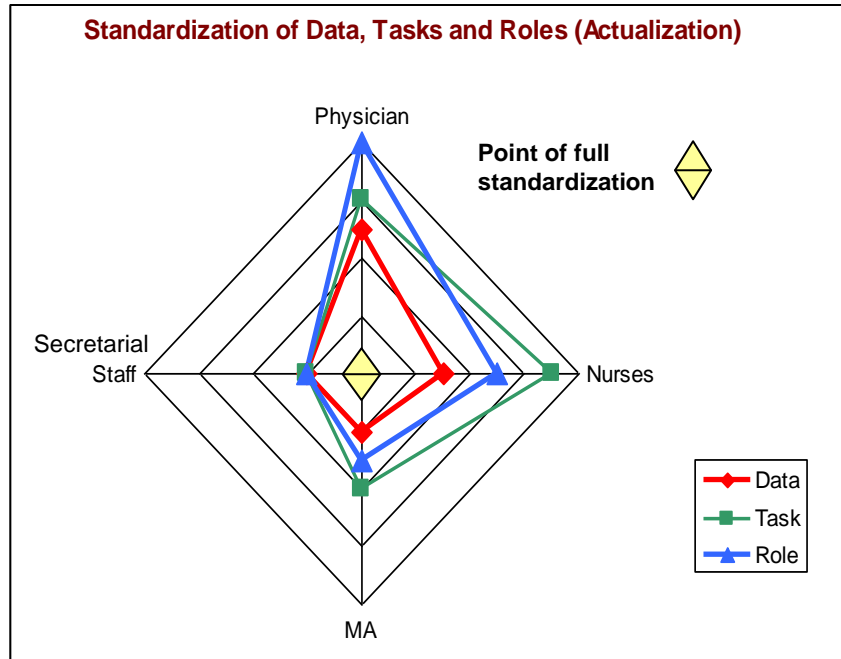
Figure 1.4 Design Tool Describing the Process State of Standardization

Figure 1.4 also demonstrates the extent of standardization as a function of role. As expected, the level of standardization of role, tasks and data is much greater for secretarial staff than for physicians mostly due to more routine and less knowledge intensive information exchanges between staff and the patient. We have also found that within the role of a physician there are conditions when routine tasks and information exchanges are candidates for standardization. There are also circumstances that require more knowledge intensive tasks by medical assistants and nurses, which then must be supported by more flexibility. While the indicator in Figure 1.4 represents aggregates of data and tasks per role, it does not capture the distribution of tasks at the individual level, some of which are conducive to standardization while others require a more knowledge intensive focus.

Finally, Figure 1.4 conveys the idea of relationships among roles. That is, the design choices for one role directly affect other roles involved in coordinating care. For example, standardized order entry by the physician in the exam room has directly affected the tasks and role of

the secretarial staff. Also, the standardization of the MA role limits their flexibility to comply

with physician requests for non-standard tasks. Figure 1.4 also implies relationships among data,

tasks and roles; standardizing one suggests standardizing another. For example, standardizing

the MA role requires some level of standardization of their tasks and the data they use.

Patterns of the process standardization state emerged from the data across roles, clinics and

pathways of delivering care. The process design tool can be used to motivate discussion future

states of standardization across the organization, for each pathway. Using the process design

tool during the planning stages of IT and process design initiatives can highlight the

opportunities to balance standardization and flexibility and minimize unintended consequences.

1.5.2 Pathways Discussion

Pathways surfaced in the data as intuitive processes initiated by physicians to deliver care

across different levels of patient complexity and information exchanges. Pathways exist

simultaneously within primary care visits and are intuitively identified and initiated by the

physician in the exam room. While it may be possible to characterize patients as high and low

complexity prior to a visit, whether the information exchange is routine or knowledge intensive

is not always predictable and may be both unexpected and unanticipated. We discovered that

EHR embedded structure can enable and constrain the specific needs of these pathways leading

to physician and staff frustration in their quest to deliver high quality care efficiently and

effectively. Designing for the careful balance between standardization and flexibility for each

pathway is possible using the EHR embedded structure and process design initiatives.

There are several benefits to recognizing and planning for these pathways in the upfront

design of a process redesign or IT implementation. First, designing for these pathways can

minimize unintended consequences as shown in our dynamics model (Figure 1.2), as well as

minimize the number of iterations to reach an acceptable process state of standardization.   At

Medical Clinic, tension between staff members arose when poor design for information

exchanges caused inefficiencies.   Information exchanges with the physician migrated from

scribbled notes and verbal interruptions to electronic messages.   A physician commented on the

inefficiency of using electronic messages for all communication needs:

> I don't think staff realizes I'm not constantly sitting on my computer and reading my
> messages. They feel as soon as they put that message somehow the light blinks in my
> brain and I've seen that message, because sometime they can put a little red light, or
> the little red sign, that means urgent message. If I'm doing a physical for half an
> hour, and then I come out and I see that the patient is in other room, I'm just walking
> in there to see patient. So it could be 45 minutes, and this message could say,
> patient's going to be there for the next 15 minutes, or she wants to be called, if she
> doesn't get a call she's going to ER. Before my nurse would walk over to me. Mrs.
> Smith wants to talk to you right now, should she go to ER, or do you want to see her?
> Now there's a little red, little message in my inbox somewhere which I'm not
> looking. So how often am I supposed to go back to my office and look at my
> messages? (PCP, site J, Round 3)

Secondly, if the pathways are not clearly recognized, designed for and communicated, then

physicians have to figure out how to process these intuitive pathways on their own.   Without

direction, individual users make decisions based on the conditions identified in Figure 1.2 and

the individual goals to achieve high quality patient care.   These actions taken collectively can

influence the ability of an organization to realize its goals.

Another benefit of designing for pathways upfront is maximizing the allotted time for

learning the system and minimizing the iterations in design after the implementation.   One nurse

commented on the desire to learn more and the frustration of system changes:

> And I'm just so busy that I don't even have time. You know, it would probably
> behoove me to sit down and do them, but I know that there are a lot of nurses that do
> use them here. Because I'll look at say what's that? Why don't I have that? And
> they say because you didn't set it up. And I'll say oh, jeez. So I think it's beneficial.
> I think I could very much benefit from it, but I'm not currently using it to its optimum
> ability… We were learning a lot of different ways and getting trained one way, and

then someone decides that it's time to change that and it no longer works. And it got changed in midstream, and that was a drawback because we weren't really prepared for that properly. I'd be putting in something for six months one way, and then one day it doesn't work. (Nurse, site H, R3)

Prior to EHR implementation at Medical Clinic, pathways existed and were highly influenced by the conditions of individual style of the physicians and complexity of the patient. As a future study, the four pathways could be expanded to include individual style as a third dimension. For instance, each pathway might reflect different choices based on the individual's level of computer experience. The value of designing for individual style is shared by a physician:

I went to medical school, a person was not required to take any typing lesson, so that's really, yeah, I could go take typing lessons and do online, whatever, but I don't have that kind of time. When I'm working 12 hours a day, plus that housework that I have to do, I'm really dead by the time I go home, I can't sit on the computer and learn to type, which I have the program. (PCP, site J, Round 2)

At Medical Clinic, the EHR system imposed embedded structures conducive but not perfectly matched to a single more-standardized pathway, resulting in many unintended consequences. Through the actions of different influencers at Medical Clinic, the number (Figure 1.3) and shape (Figure 1.4) of pathways changed over time, adapting to both patient complexity and the type of information exchange. In the future, as conditions and consequences shape the process standardization state of the pathways, actions will arise in response as shown in our model (Figure 1.2). Consequently, it is quite possible for the process states described by the design tool in Figure 1.4 to take on new shapes. The number of unique pathways may also increase or decrease based on dynamic changes in conditions and the actions of the influencers. As health care evolves to a more patient-centered care model, similar to the Chronic Care Model (Wagner, 1998) or Patient-centered Medical Home (Rogers, 2008; Rosenthal, 2008), the number and type of the pathways could be completely different. One question is whether EHR systems

will have functionalities to support these pathways, leading to productive interactions between the informed activated patient and the prepared, proactive practice team in the microsystem (Wagner, 1998; Nelson et al., 2007).

1.6 Conclusions

In the current health care environment, new ailments, medical regimens and technologies emerge daily. Additionally, the presence of chronic diseases and their related comorbid conditions are rising and constitute over half of all primary care visits (IOM, 2001). Patient complexity and preferences are consequently escalating a more knowledge intensive environment for primary care which physicians must manage. During a primary care visit, physicians encounter different paths for processing patient information, some of which can be standardized and some of which require a more knowledge intensive and flexible path. The structure embedded in the EHR system can standardize data and tasks in these paths in a way that might not be conducive to all patient complexities and provider styles. A health care organization can benefit from identifying pathways that are conducive to standardization and those which are more conducive to flexibility. By employing a process lens in the context of primary care, health care organizations are better equipped to recognize and then alleviate tension caused by the structure of an EHR system.

Using a grounded theory methodology, we developed a model explaining the overall dynamics between the conditions, outcomes, and actions of a process state relative to standardization (Figure 1.2). We also examined how the embedded elements of an EHR can support different levels of standardization for different pathways. Four pathways emerged in response to the EHR intervention and over time became more defined as influencers took action to alleviate unintentional consequences and interrelated conditions. The results of this study

showed that the complexity of the patient (type of patient, type of visit and preferences) and the information exchange (routine or knowledge intensive) affect how much a primary care visit or process (data, tasks, and roles) can be standardized. The intersection of these conditions created the four distinct pathways: standardization, multiple standardized paths, structure with flexibility and flexibility (Figure 1.3).

The conditions of physician style, site characteristics and the EHR embedded structures intersect and shape each pathway and are levers for decision makers to consider when modifying the pathways.   Additionally, we created a process design tool (Figure 1.4) to qualitatively describe the current and future state of standardization in each pathway or across the organization for process improvement planning or intervention analysis. We offered specific suggestions for health care organizations to consider when making decisions regarding IT and process design for each pathway. In the absence of upfront design, individuals make choices about how and to what extent to use the EHR embedded structures to deliver patient care. They create tasks and information exchanges aligned with their individual style and goals. These actions taken collectively can influence the ability of an organization to realize its goals. HC organizations can benefit by designing for specific pathways before EHR implementation to minimize the tension from a one-size-fits-all approach, to shorten the time to return to pre-EHR efficiencies, to minimize the probability of a failed EHR implementation, and to maintain patient and staff satisfaction.

There are some limitations to our research. We studied one EHR system at one health care organization, although we considered multiple sites within the organization. Additional longitudinal research should test the viability of the process state model as well as designs that emerge around the distinct pathways for different EHR platforms and for different organizational

and societal contexts. Additionally, our model and design for pathways assume that individual style is not a strong condition shaping pathways. Future work could explore the possibility of characterizing individual style as a third dimension in defining pathways for IT and process design initiatives. Properties like computer expertise are specific to individuals, and might be considered as separate pathways within the current pathways framework. Incorporating individual style could surface opportunities to further align the work of the individual users delivering patient care within the goals of the organization.

Our study offers health care organizations a framework for thinking about the extent to which standardization should be promoted in a patient-centered primary care process. Our study contributes to a better understanding of the process state of standardization and how medical management should think about, and make decisions about, EHR embedded structures and process design in primary care. At what point does standardization limit the ability to capture a patient's story or limit the ability for a provider to create the most efficient and effective patient plan?

1.7 Acknowledgments

1.8 Appendix A Definition of Standardization

Standardization of Data

Standardization of data refers to standardizing the data fields that should be entered (for example: height, weight, and demographics); the data values that can be entered by requiring selection from lookup tables for diagnosis and order entry (for example: prescriptions, labs and pharmacy locations); and the form and content of data outputs (for example: letters, forms and patient visit summaries). We organize our data standardization results into the properties of information use, information availability and data quality.

Information use was increased by Medical's EHR through its shorthand methods for documenting single words, phrases or texts; uploading patient specific data into a form; selecting common terms from pull down lists; and utilizing standardized templates for visit notes, patient instructions or telephone encounters. As a practice manager noted,

> Some of the dot phrases I just love. When you're putting in a message and you hit dot-PH, and the person's phone number populates right into your note without having to look it up or write it down, you just verify it with the patient. (Practice Manager, site C, round 1)

One physician commented on the ability to use the information in the system collectively to deliver better care to patients, noting:

> If I'm seeing a diabetic, you click on the diabetes smart set, and it gives you all of their relevant numbers regarding the diabetes. Or if they're anemic, it just, you can bring down all of the relevant lab work to that. So that helps in managing our chronic illnesses.
> Q: So you find those useful then.
> A: Yes, very useful.
> Q: OK. Over all does it help you deliver better care?
> A: Yeah, it does. Yeah. (PCP, site F, round 3)

The information in the system is also being used by management for problem solving and performance analysis. One practice manager commented,

"It's made problem solving a lot easier in my job. It's made not just problem solving like a patient calls and says, this is the problem, or I talked to somebody yesterday, and I don't know where things are at. So for me to be able to jump in from a management point of view and say, OK, well let me see what's going on, I can fix that much easier. I can resolve those issues while the person's on the phone, as opposed to having to pull the chart and review the notes and try to figure out who said this and who signed. You know what? Much, much easier. So yeah, problem solving, identifying what's what. But it's also a watch, I don't mean this in a negative way, it's really helped me identify where staff's growth areas are. (Practice manager, site J, round 1)

Data quality was increased by Medical's EHR by ensuring that the data entered into or selected within the EHR are complete and accurate. Shorthand phrases to populate letters, drop down options for medications and pharmacy locations and safeguards on data entry of height and weight are examples of data standardization mechanisms enabling data accuracy and complete-ness. Interviewees noted that there is no guess work in interpreting instructions or messages in the EHR, thereby minimizing rework for clarification and frustration when information is missing. As a secretarial staff member (site C, round 1) noted, it's "not a question anymore about what the doctor wants because he has to put it in the system."

Data availability was increased by Medical's EHR by having the ability to extract patient information from the EHR in a time efficient manner versus waiting for a paper chart. As a nurse stated,

I think it's wonderful because everything I need is right there. I don't have to leave that exam room. I don't have to go call for charts. I can find what I'm looking for from appointments to when their last EKG was six years ago. So it is such a huge, huge difference. (Nurse, site C, round 3)

However, some physicians believed there was less information available because some data did not fit well into standardized fields or could not be entered through shortcuts. This was compounded by the incomplete and in some cases inaccurate initial data load of historical patient information into the EHR. Nurses and doctors have also developed paper cheat sheets that they store near their workstations as a way of minimizing the time spent searching for specific data.

Standardization of Tasks

Standardization of tasks refers to workflows and processes, including what tasks are done, whether the right tasks are done, and whether tasks are done in the same way. The embedded scripts in the EHR software, which dictate the tasks of medical staff and in some cases the sequence of tasks and workflow, strongly influence the extent to which the ***process is defined***. One nurse from site C during round 2 noted, "I can do like 50 refills in minutes by the process."

Also, some physicians mentioned a benefit that the system explicitly tracks the status of their tasks, contributing to physician satisfaction. One physician at site A added,

> It has made very explicit what needs to be done, but probably the biggest single change is that when the messages come in, it's very clear on the screen there that you have prescriptions to fill. You have patient calls to answer. You know exactly what you have to do, and whether it's been accomplished. There are no lost pieces of paper. (PCP, site A, round 2)

The ability to standardize tasks is also a function of the level of ***process variability*** and complexity in those tasks outside of the EHR system. One physician explains "My patients are old, multi-problem, so [templates] don't work." (PCP, site A, round 2)

The system design of tasks in some cases has imposed undesirable changes in work routines arising from the rigid ways each task can be performed in the system. Some nurses believe that working through the triage protocol takes more time than using free text for the same information especially when there is variability in how each patient articulates the reason for the call. One nurse stated,

> The problem is triaging a patient is not scripted. I'm not going to sit there and ask all of those questions and get those answers. You can get them from a story. You can get them from their conversation, and it doesn't come in the order of, so [templates] are not good for triaging. (Nurse, site C, round 2)

Standardization of Roles

Standardization of roles is associated with increased substitutability between medical staff members and increased use of best practices for the delivery of care, e.g., the use of physician protocols for chronic disease management and reminders for preventative care. EHR systems offer features like pools and tools to define roles by, for example, offering choices in levels of permission to access data and perform tasks.

The EHR system imposes embedded routines on the physician role in terms of what tasks must be performed to complete a visit and the type of information exchanged between roles. For example, physicians are now booking appointments, and ordering medications and labs in the exam room. This has caused a concern amongst physicians as to who should be doing what and a concern among other staff members about job security as tasks have shifted. ***Bucket of Tasks*** is a dimension of role standardization that informants discussed passionately. One primary care physician commented, "So I'm doing a lot of work that down the stream people don't have to do." (PCP, site C, phase 3)

Role identity is another dimension in the design of role standardization. The tasks and skills required within jobs have evolved to become more computer-based and in some cases less desirable. Specifically, the amount of computer use and typing tasks have increased considerably with the implementation of the EHR system leaving some medical staff feeling a sense of role mismatch. One nurse from site A, round 2 noted I'm "glued to a chair, staring at a computer, and it is dissatisfying. I love to be on my feet and using my hands and being clinical."

Physicians and nurses also complained about the lack of interaction with the other medical staff and patients. One physician from site C, round 3 commented,

> I feel that we're spending an awful lot of time looking at that computer, setting up dates of return and labs and things, that, and I'm not doing the patient care.

1.8 Appendix B Pathway Examples

Table 1.5 Pathway Examples on Outcomes: Matrix of conditions/consequences

Pathway	Outcomes
Standardization	**PCP (site H, round 3):** "we feel it's a better outcome with [the EHR] because when we do the [templates], everything is there. So the labs, pertaining labs, pertaining testing, everything is there. So you don't miss things. It's all in one order. So actually that also helps with the patient care. It's reminding the physicians, OK, you didn't do this. So at least you can think about, OK, I need to order this. I think that's another good feature." **Nurse (site F, round 3):** "no longer do they have to see a UTI patient. We take care of that issue. All they have to do is sign it, and send it once we get the information. They no longer have to look anything up. If I'm already looking the lab up, why don't I copy, and paste it, and put it on my telephone message? And then they can see everything right in the message. So we have made it a lot easier for them." **PCP (site A, round 2):** "Messages from nurses are better. It has forced them to ask standardized questions; with responses taken down in an order...My instructions back to the nurses are better." **PCP (site G, round 2):** "The first time I ordered, tried to order an Aerochamber, a list of about 25 options came up. Now, I honestly didn't know there were that many Aerochambers in the world. And I couldn't pick out the one I wanted." **PCP (site I, round 1):** "I think eventually that would be a reasonable goal that say, the patient prep is more consistent and standardized because they have to go through the [templates] in [the EHR]. But, as you know, it's to get any large group to agree on certain processes is hard, and I think what we need to really focus on is the outcomes, and I think the processes will then fall in place." **MA (site A, round 3):** "Lately, we've been doing more background. It's a brand-new process. We didn't before have to look at previous Pap smears or, you know, previous biopsies or the history, and now we do. It makes for a more accurate current Pap smear, but it wasn't implemented before. It's brand-new." **PCP (site H, round 3):** "but one thing is you kind of stop thinking because everything is there. Before we had to memorize our antibiotic dose, and the thing. Now it's there, you can, find. You can find everyone's TID."
Multiple Standardized Paths	**PCP (site F, round 1):** "I think one thing that looks like it will be really helpful is they have these flow charts so if you have a diabetic patient, you click on the diabetes flow chart and automatically it brings up all their blood sugars, their hemoglobin, A1C's, their cholesterols." **PCP (site D, round 3):** "and I see a lot of chronic disease, so I have actually learned how to bend the system to that type of patient. So for example, if you come to see me and it's just a simple high blood pressure, you have high blood pressure, see you once or twice a year whatever. I now actually have my notes that I have formatted for that visit, so I pull out my note, pops in there, and there's a few questions that I have to answer, a few things I have to change. So I've bent the system to that, so that makes things very, very easy." **Practice Manager (site E, round 3):** As the MA puts in, there's certain things that they have to, you know do to make sure that it's either the pulse ox or the inhaler to make sure that their peak flow is good. All those same things are standardized, so that has been very good because we used to have that all over the board. Some people would do it. Some people wouldn't do it. Some people required it. Some people didn't require it. So that has definitely, you know, those questions come up, have you done this? Have you done that? So that is good. So that piece, I think, is better."
Structure with Flexibility	**PCP (site A, round 3):** "With the old system, you could not be sure that what went from your brain to the note actually went to an order somewhere. With this system, you can't close unless you actually see the orders, or at least it will tell you, there are no orders. So, if you are intending to do a blood test or an x-ray, or a referral or something, and you look down and see it's not there before you close it, you have an opportunity to do that. So, it's much better from that standpoint." **PCP (site C, round 2):** "I'm not doing the patient care. Then going out and trying to use the system the Smart Sets which, you're using three or four Smart Sets if you get somebody in with different things. And then when you try to read them afterwards, they don't make sense. There's not a flow, there's not a, I could read the note, and I understood what I saw the patient for. It's kind of scattered, and there's almost too much that you have to go through. I still don't feel comfortable with it, and so I'm not a fan." **MA (site A, round 3):** "It's if you put in a blood pressure that's over 140 over 80; it stops you and asks you is this correct? And it's very rare that somebody, you know, we have a lot of patients with hypertension so, you know, it's just that redundancy. But it is a precaution and it is making you

	double-check yourself. So it's looking out for human error so I guess it's, it does it with the weight also. If you put in a weight that's over a certain amount, a percentage of their last visit, it alerts you, which, again, it's looking for human error, you know, which can happen. So it's kind of a love/hate kind of thing because it does slow you down and you need to get out, but you need to be accurate at the same time." **PCP (site J, round 3):** "they don't tell you that in the beginning. I'm here for the back pain, so you put the back pain, you open the Smart Set for back pain, and you do the back pain thing. And then say, and what about my rash? And remember my heartburn, and you know my migraines are really not getting that better anymore. So now you really should have put in the back pain and the headaches, and the rash, and UTI, or whatever, so then you have like a 20 Smart set and although you can open only one Smart set at a time. So then you click to the back pain Smart set, you got it out then go back, and then open the rash Smart set and then you order the skin test and whatever. So it takes you longer. And I don't like the progress note still, because the progress note for the back pain will take me a much longer time because they have to have all the categories of back pain. It's not this was rather simple, she was shoveling the snow, there's nothing on exam, but still to document that on the [EHR] note will take me much longer, so I'm better off dictating, so I'm still dictating."
Flexibility	**PCP (site J, round 2):** "I just get up and walk and talk to them. I said, listen, I'm not going to write nothing because I can't type that fast. So sometime, it's better for me to just walk and say, tell patient, duh, duh, duh, duh, duh, and just document it, I'm not typing." **PCP (site A, round 3):** "It's not state-of-the-art and not where I think it needs to be in terms of trying to make the process easier for the docs, but it's unquestionably improved quality of care. I would be very unambiguous about that. Because the information about the patient and the process of care for the patient is available, legible, retrievable, sharable. I mean, you know, everything that you would want in terms of somebody coming in and being able to see what was done, what tests were done, what results were there, what drugs were prescribed, were they picked up, laboratory results and things like that. It is not uniformly, though, retrievable in terms of the process of thinking because people are very inconsistent about how they do the notes. If you do a canned note from this system and just fulfill your obligations, you'll never get a flavor for what actually happened to the patient within the interview and the physical exam. If you dictate then you're kind of defeating the purpose of the system, but that remains the best way to get a note that actually reflects nuances of how the patient looked, what the physician who was caring for them was thinking, why they chose to do what they did. So that piece is not uniformly available. If I want to know how your CT scan came out, this is beautiful. If I want to know if you had a colonoscopy and, if so, when and what did they find, perfectly fine, last mammogram, what was on it, last CT scan of the chest, all of that stuff absolutely fine, but not the process of op. But what I like least about it is that there's still too much that's canned and there's not enough support for people who are trying to make it not canned." **PCP (site C, round 3):** "it depends on whether somebody is using the Smart Sets and if they use the Smart Sets, it is very difficult to go through those. Because they are using a smart set for each thing and so it is hard to necessarily pick apart things if they've used umpteen Smart Sets. You also have to roll through things that you're saying, I don't care to see that, you know, what did they do. If somebody is dictating, I think it's much easier to figure out what they're doing. If they're using the language of the Smart Sets, sometimes it makes no sense. I mean it just isn't easy, because they're, they're doing paper cutter type of, type of thing, and you're saying, did they punch that in because it was just one of the things to punch in or was that important for them. So and it varies from individual to individual as to how they do it. **PCP (site H, round 3):** Each patient is unique. Especially with a complicated, difficult patient, it's individual. So it's hard to do common notes for those patients. Maybe if we could implement dragon software with this, so we don't have to dictate and then wait for a dictation for a few days. I think that would be a good option. **Practice Manager (site C, round 3):** The goal in my mind is to get the docs on [real-time transcription speech recognition software]. Because I'm not going to force somebody to type if they can't type, and you're not going to force them anyways. So you need to find something that they're comfortable with that can increase their efficiencies. **MA (site D, round 3):** So that's why it's nice for the scribe to be in there? A: because they can be the one focusing on the computer while the doctor's still focusing on the patient. Because like I said, if you miss one little category it won't let you continue and then you're looking back to try to figure out what's wrong, why isn't it letting me move on? And that puts the doctor behind, a simple little thing like that, because then they start getting frustrated because the patient's still going, and they're trying to concentrate on them but they're also trying to concentrate well why isn't my computer working for me? So that's where I come in and he doesn't have to worry about all of that.

1.9 References

Apter, A. J., M. K. Paasche-Orlow, J. T. Remillard, I. M. Bennett, E. P. Ben-Joseph, R. M. Batista, J. Hyde, and R. E. Rudd. "Numeracy and Communication with Patients: They are Counting on Us," *Journal of General Internal Medicine* (23:12), Dec, 2008, pp. 2117-2124.

Argote, L. "Input Uncertainty and Organizational Coordination in Hospital Emergency Units," *Administrative Science Quarterly* (27:3), 1982, pp. 420-434.

Barua, A., S. Ravindran, and A. Whinston. "Enabling Information Sharing within Organizations," *Information Technology & Management* (8:1), 01, 2007, pp. 31-45.

Bates, D. W., D. L. Boyle, and J. M. Teich. "Impact of Computerized Physician Order Entry on Physician Time," *Proceedings / the ...Annual Symposium on Computer Application [Sic] in Medical Care.Symposium on Computer Applications in Medical Care* 1994, pp. 996.

Bates, D. W., M. Cohen, L. L. Leape, J. M. Overhage, M. M. Shabot, and T. Sheridan. "Reducing the Frequency of Errors in Medicine using Information Technology," *Journal of the American Medical Informatics Association : JAMIA* (8:4), Jul-Aug, 2001, pp. 299-308.

Bates, D. W., G. J. Kuperman, S. Wang, T. Gandhi, A. Kittler, L. Volk, C. Spurr, R. Khorasani, M. Tanasijevic, and B. Middleton. "Ten Commandments for Effective Clinical Decision Support: Making the Practice of Evidence-Based Medicine a Reality," *Journal of the American Medical Informatics Association : JAMIA* (10:6), Nov-Dec, 2003, pp. 523-530.

Bates, D. W., M. Ebell, E. Gotlieb, J. Zapp, and H. C. Mullins. "A Proposal for Electronic Medical Records in U.S. Primary Care," *Journal of the American Medical Informatics Association* (10:1), January 1, 2003, pp. 1-10.

Berner, E. S., D. E. Detmer, and D. Simborg. "Will the Wave Finally Break? A Brief View of the Adoption of Electronic Medical Records in the United States," *Journal of the American Medical Informatics Association : JAMIA* (12:1), Jan-Feb 2005, pp. 3-7.

Berwick, D. M. "A User's Manual for the IOM's 'Quality Chasm' Report," *Health Affairs* (21:3), May 1, 2002, pp. 80-90.

Bodenheimer, T., E. H. Wagner, and K. Grumbach. "Improving Primary Care for Patients with Chronic Illness," *JAMA : The Journal of the American Medical Association* (288:14), Oct 9, 2002, pp. 1775-1779.

Brown, J. S., and P. Duguid. "Balancing Act: How to Capture Knowledge without Killing it," *Harvard Business Review* (78:3), 2000, pp. 73-80.

Bruce, M., L. Daly, and K. B. Kahn. "Delineating Design Factors that Influence the Global Product Launch Process*," *Journal of Product Innovation Management* (24:5), 09/24, 2007, pp. 456-470.

Carayon, P., A. Schoofs Hundt, B. Karsh, A. P. Gurses, C. J. Alvarado, M. Smith, and P. Flatley Brennan. "Work System Design for Patient Safety: The SEIPS Model," *Quality and Safety in Health Care* (15:suppl_1), December 1, 2006, pp. i50-58.

Chaudhry, B., J. Wang, S. Wu, M. Maglione, W. Mojica, E. Roth, S. C. Morton, and P. G. Shekelle. "Systematic Review: Impact of Health Information Technology on Quality, Efficiency, and Costs of Medical Care," *Annals of Internal Medicine; Annals of Internal Medicine* (144:10), May 16, 2006, pp. 742-752.

Cherry, D. K., E. Hing, D. A. Woodwell, and E. A. Rechtsteiner. "National Ambulatory Medical Care Survey: 2006 Summary," *National Health Statistics Reports* (3), Aug 6, 2008, pp. 1-39.

Crosson, J. C., C. Stroebel, J. G. Scott, B. Stello, and B. F. Crabtree. "Implementing an Electronic Medical Record in a Family Medicine Practice: Communication, Decision Making, and Conflict," *Annals of Family Medicine* (3:4), Jul-Aug, 2005, pp. 307-311.

Davenport, T. H., J. G. Harris, and S. Cantrell. "Enterprise Systems and Ongoing Process Change," *Business Process Management Journal* (10:1), February 2004, pp. 16-26.

Davidson, E. J., and W. G. Chismar. "The Interaction of Institutionally Triggered and Technology-Triggered Social Structure Change: An Investigation of Computerized Physician Order Entry," *MIS Quarterly* (31:4), 12, 2007, pp. 739-758.

Davidson, E., and W. G. Chismar. "Planning and Managing Computerized Order Entry: A Case Study of IT-Enabled Organizational Transformation," *Topics in Health Information Management* (19:4), 1999, pp. 47-61.

Dickson, E. W., Z. Anguelov, D. Vetterick, A. Eller, and S. Singh. "Use of Lean in the Emergency Department: A Case Series of 4 Hospitals," *Annals of Emergency Medicine* (54:4), Oct, 2009, pp. 504-510.

Fischer, G., M. D. Fetters, A. P. Munro, and E. B. Goldman. "Adverse Events in Primary Care Identified from a Risk-Management Database," *The Journal of Family Practice* (45:1), Jul, 1997, pp. 40-46.

Frankel, R., A. Altschuler, S. George, J. Kinsman, H. Jimison, N. R. Robertson, and J. Hsu. "Effects of Exam-Room Computing on Clinician-Patient Communication: A Longitudinal Qualitative Study," *Journal of General Internal Medicine : Official Journal of the Society for Research and Education in Primary Care Internal Medicine* (20:8), August 2005, pp. 677-682.

Frieling, W. "Beyond 'Meaningful use'. Regional Health Information Exchanges just as Important to Healthcare IT," *Modern Healthcare* (39:32), Aug 10, 2009, pp. 22.

Gittell, J. H., R. Seidner, and J. Wimbush. "A Relational Model of how High-Performance Work Systems Work," *Organization Science,* July 21, 2009, pp. 1-17.

Glaser, B. G., and A. L. Strauss. *The Discovery of Grounded Theory: Strategies for Qualitative Research*, Aldine Pub. Co, Chicago, 1967.

Goldschmidt, P. C. "HIT and MIS: Implications of Health Information Technology and Medical Information Systems," *Communications of the ACM* (48:10), October 2005, pp. 69-74.

Green, C. J., P. Fortin, M. Maclure, A. Macgregor, and S. Robinson. "Information System Support as a Critical Success Factor for Chronic Disease Management: Necessary but Not Sufficient," *Int J Med Inform.* (75:12), Aug 16, 2006, pp. 818-828.

Hall, J. M., and M. E. Johnson. "When should a Process be Art, Not Science?" *Harvard Business Review* (87:3), 03, 2009, pp. 58-65.

Helm, S.,2nd, and H. C. Hansen. "Information Technology in the Interventional Pain Practice: Electronic Medical Records, Practice Management Software, and Document Management," *Pain Physician* (7:3), Jul, 2004, pp. 357-364.

Hing, E. S., C. W. Burt, and D. A. Woodwell. "Electronic Medical Record use by Office-Based Physicians and their Practices: United States, 2006," *Advance Data* (393), Oct 26, 2007, pp. 1-7.

Holden, R. J., and B. Karsh. "The Technology Acceptance Model: Its Past and its Future in Health Care," *Journal of Biomedical Informatics* (43:1), 2, 2010, pp. 159-172.

Imai, M. *Kaizen (Ky'Zen), the Key to Japan's Competitive Success*, Random House Business Division, New York, 1986.

Institute of Medicine (U.S.). Committee on Quality of Health Care in America. *Crossing the Quality Chasm : A New Health System for the 21st Century*, National Academy Press, Washington, D.C., 2001.

Kaplan, B. "Evaluating Informatics Applications - Clinical Decision Support Systems Literature Review," *International Journal of Medical Informatics* (64:1), 2001, pp. 15-37.

Karsh, B. "Beyond Usability for Patient Safety: Designing Effective Technology Implementation Systems," *British Medical Journal: Quality and Safety in Healthcare* (13:5), 2004, pp. 388-394.

Lamb, R., and R. Kling. "Reconceptualizing Users as Social Actors in Information Systems Research," *MIS Quarterly* (27:2), 2003, pp. 197-223.

Liker, J. K. *The Toyota Way : 14 Management Principles from the World's Greatest Manufacturer*, McGraw-Hill, New York, 2004.

Likourezos, A., D. B. Chalfin, D. G. Murphy, B. Sommer, K. Darcy, and S. J. Davidson. "Physician and Nurse Satisfaction with an Electronic Medical Record System," *The Journal of Emergency Medicine* (27:4), November 2004, pp. 419-424.

Linder, J. A., J. Ma, D. W. Bates, B. Middleton, and R. S. Stafford. "Electronic Health Record use and the Quality of Ambulatory Care in the United States." *Archives of Internal Medicine* (167:13), 2007, pp. 1400-1405.

Litvak, E., P. I. Buerhaus, F. Davidoff, M. C. Long, M. L. McManus, and D. M. Berwick. "Managing Unnecessary Variability in Patient Demand to Reduce Nursing Stress and Improve Patient Safety," *Joint Commission Journal on Quality and Patient Safety / Joint Commission Resources* (31:6), June 2005, pp. 330-338.

Makoul, G., R. H. Curry, and P. C. Tang. "The use of Electronic Medical Records: Communication Patterns in Outpatient Encounters," *Journal of the American Medical Informatics Association : JAMIA* (8:6), Nov-Dec 2001, pp. 610-615.

Maxwell, M. "EMR: Successful Productivity Tool for Modern Practice," *Health Management Technology* (20:9), October 1999, pp. 48-49.

McGlynn, E. A., S. M. Asch, J. Adams, J. Keesey, J. Hicks, A. DeCristofaro, and E. A. Kerr. "The Quality of Health Care Delivered to Adults in the United States," *The New England Journal of Medicine* (348:26), Jun 26, 2003, pp. 2635-2645.

Murray, M. "Strategies for the Successful Implementation of Workflow Systems within Healthcare: A Cross Case Comparison," *System Sciences, 2003.Proceedings of the 36th Annual Hawaii International Conference on*2003, pp. 10.

National Committee for Quality Assurance. "HEDIS 2007 Technical Specifications for Physician Measurement." 2007.

Nelson, E. C., P. B. Batalden, and M. M. Godfrey. *Quality by Design : A Clinical Microsystems Approach*, Center for the Evaluative Clinical Sciences at Dartmouth; Jossey-Bass/Wiley, Lebanon, NH; San Francisco, 2007.

Nemeth, L. S., C. Feifer, G. W. Stuart, and S. M. Ornstein. "Implementing Change in Primary Care Practices using Electronic Medical Records: A Conceptual Framework," *Implement Sci.* (3:1), Jan 16, 2008, pp. 3.

Nutting, P. A., W. L. Miller, B. F. Crabtree, C. R. Jaen, E. E. Stewart, and K. C. Stange. "Initial Lessons from the First National Demonstration Project on Practice Transformation to a Patient-Centered Medical Home," *Annals of Family Medicine* (7:3), May-Jun, 2009, pp. 254-260.

Orlikowski, W. J., J. Yates, K. Okamura, and M. Fujimoto. "Shaping Electronic Communication: The Metastructuring of Technology in the Context of use," *Organization Science* (6:4), /07Jul/Aug95, 1995, pp. 423-444.

Ovretveit, J., T. Scott, T. G. Rundall, S. M. Shortell, and M. Brommels. "Improving Quality through Effective Implementation of Information Technology in Healthcare," *International Journal for Quality in Health Care : Journal of the International Society for Quality in Health Care / ISQua* (19:5), Oct, 2007, pp. 259-266.

Parks, C. M. "The Bare Necessities of LEAN," *Industrial Engineer: IE* (35:8), 08, 2003, pp. 39.

Pizziferri, L., A. F. Kittler, L. A. Volk, M. M. Honour, S. Gupta, S. Wang, T. Wang, M. Lippincott, Q. Li, and D. W. Bates. "Primary Care Physician Time Utilization before and After Implementation of an Electronic Health Record: A Time-Motion Study," *Journal of Biomedical Informatics; Journal of Biomedical Informatics* (38:3), June 2005, pp. 176-188.

Poissant, L., J. Pereira, R. Tamblyn, and Y. Kawasumi. "The Impact of Electronic Health Records on Time Efficiency of Physicians and Nurses: A Systematic Review," *Journal of the American Medical Informatics Association : JAMIA* (12:5), Sep-Oct, 2005, pp. 505-516.

Pollock, N., R. Williams, and R. Procter. "Fitting Standard Software Packages to Non-Standard Organizations: The " Biography' of an Enterprise-Wide System," *Technology Analysis & Strategic Management* (15:3), 09, 2003, pp. 317.

Poon, E. G., T. K. Gandhi, T. D. Sequist, H. J. Murff, A. S. Karson, and D. W. Bates. ""I Wish I had seen this Test Result Earlier!": Dissatisfaction with Test Result Management Systems in Primary Care," *Archives of Internal Medicine* (164:20), Nov 8, 2004, pp. 2223-2228.

Poon, E. G., A. K. Jha, M. Christino, M. M. Honour, R. Fernandopulle, B. Middleton, J. Newhouse, L. Leape, D. W. Bates, D. Blumenthal, and R. Kaushal. "Assessing the Level of Healthcare Information Technology Adoption in the United States: A Snapshot," *BMC Medical Informatics and Decision Making [Computer File]; BMC Medical Informatics and Decision Making [Electronic Resource]*, Jan 5, 2006, pp. 1.

Pope, C., and N. Mays. *Qualitative Research in Health Care*, Blackwell Pub. BMJ Books, Malden, Mass., 2006.

Poulymenopoulou, M., F. Malamateniou, and G. Vassilacopoulos. "Specifying Workflow Process Requirements for an Emergency Medical Service," *Journal of Medical Systems* (27:4), 2003, 2003, pp. 325-335.

Reid, P. P., W. D. Compton, J. H. Grossman, G. Fanjiang, National Academy of Engineering, Institute of Medicine, and National Academies Press. *Building a Better Delivery System: A New Engineering/Health Care Partnership*, National Academies Press, Washington, D.C., 2005.

Rogers, J. C. "The Patient-Centered Medical Home Movement--Promise and Peril for Family Medicine." *Journal of the American Board of Family Medicine (Formerly Journal of the American Board of Family Practice)* (21:5), 2008, pp. 370.

Rosenthal, T. C. "The Medical Home: Growing Evidence to Support a New Approach to Primary Care." *Journal of the American Board of Family Medicine (Formerly Journal of the American Board of Family Practice)* (21:5), 2008, pp. 427.

Scott, J. T., T. G. Rundall, T. M. Vogt, and J. Hsu. "Kaiser Permanente's Experience of Implementing an Electronic Medical Record: A Qualitative Study," *BMJ (Clinical Research Ed.)* (331:7528), December 3, 2005, December 3, 2005, pp. 1313-1316.

Sharrock, R. "Road to Gold," *Industrial Engineer: IE* (39:5), 05, 2007, pp. 44-48.

Stalk, G., P. Evans, and L. E. Shulman. "Competing on Capabilities: The New Rules of Corporate Strategy," *Harvard Business Review* (70:2), Mar/Apr92, 1992, pp. 54-66.

Strauss, A. L., and J. M. Corbin. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, Sage Publications, Thousand Oaks, 1998.

Volkoff, O., D. M. Strong, and M. B. Elmes. "Technology Embeddedness and Organizational Change," *Organization Science* (18:5), September-October 2007, pp. 832-848.

Volkoff, O., D. M. Strong, and M. B. Elmes. "Understanding Enterprise Systems-Enabled Integration," *European Journal of Information Systems* (14:2), June 2005, pp. 110-120.

Wagner, E. H. "Chronic Disease Management: What Will it Take to Improve Care for Chronic Illness?" *Effective Clinical Practice : ECP* (1:1), Aug-Sep, 1998, pp. 2-4.

Womack, J. P., and D. T. Jones. *Lean Thinking : Banish Waste and Create Wealth in Your Corporation*, Free Press, New York, 2003.

Wood, N. "Lean Thinking: What it is and what it Isn't," *Management Services* (48:2), 02, 2004, pp. 8-10.

Yackel, T. R., and P. J. Embi. "Unintended Errors with EHR-Based Result Management: A Case Series," *Journal of the American Medical Informatics Association : JAMIA* (17:1), Jan-Feb, 2010, pp. 104-107.

Zhou, L., C. S. Soran, C. A. Jenter, L. A. Volk, E. J. Orav, D. W. Bates, and S. R. Simon. "The Relationship between Electronic Health Record use and Quality of Care Over Time," *Journal of the American Medical Informatics Association* (16:4), July/August, 2009, pp. 457-464.

2.0 A Composite Model for Profiling Physicians Across Domains of Care

2.1 Introduction

Performance profiling has been recognized as a key enabler in health care reform by health policy makers and large commercial purchasers who bear the burden of supporting the world's highest health care expenditures (Kaiser Family Foundation, 2009). In an effort to bring value to their employer clients and maintain a competitive advantage, health plans have developed two types of customized programs: pay-for-performance (P4P) programs to incentify physician behavior and consumer-based decision support programs to engage patients in their health care decisions (Wodchis et al., 2007). Physician performance profiling is used to describe techniques and models for comparing physicians on quality and efficiency measures within and across sites, health care organizations and health plans. Health plans use profiling extensively in physician-oriented P4P programs to determine performance targets for contract negotiations and financial incentives at the physician and organizational levels. However, efforts to measure performance at the individual physician level face a number of challenges, including the need for sufficient sample size to support reliable profiling (Adams, 2010).

Our physician profiling model ranks primary care physician (PCP) performance across a broad range of measures using data from multiple payers and data sources. We demonstrate its use as a means to assign incentive payments at a multi-clinic health care organization, which we call Health Clinic. The model is expandable to include various domain types (i.e., acute, chronic, preventive, patient experience, and Medical Home), sources of data (i.e., Electronic Health Records, Medicaid and Medicare data) and stratification within domains for patient-mix or clinical practice characteristics. With such flexibility, health care organizations can tailor the

model to match organizational characteristics; the model was created with input from and approval by physicians at Health Clinic while incorporating the strategic mission and values of the organization. In this study, we present a methodology for building a physician profiling model incorporating criteria suggested in the literature to ensure fair comparisons across physicians. We then develop a physician profiling model at Health Clinic that addresses reliability by increasing sample sizes and variation across physicians through the use of composites, multiple payers and data sources.

2.2 Literature Review

The literature review section is organized according to two components deemed critical for framing the problem, the importance and the state of knowledge. The components are physician profiling background and the criteria for physician profiling models. We discuss each of these in turn.

2.2.1 Physician Profiling Background

Pay for performance is an attempt to address problems with the fee-for-service model (FSS) by establishing incentives for physicians to deliver care that third parties deem is necessary and appropriate to achieve the highest quality standards and best outcomes (Wodchis et al., 2007). Under the FSS model, the physicians receive payments from the patient or health plan for the services rendered, whether evidence-based, supplementary or patient demanded. FFS encourages the overuse of medicine: the more physicians do, the more money they receive. Under the P4P model, the physicians receive payments from health plans for delivering appropriate preventive care (e.g., colorectal cancer screening), chronic care (e.g., diabetes HbA1c blood sugar testing), and acute care (e.g., Pharyngitis testing for the eligible patients in their patient panel). Although pay-for-performance is typically a small add-on to fee-for-service,

a P4P program focused on a small group of measures could lead physicians toward pay-for-compliance instead of pay-for-excellence (Berwick, 2009). Rosenthal et al (2007) studied twenty-seven health plans that had P4P programs and found a lot of emphasis on efficiency and outcome measures rather than on process measures. Yet, most of the health plans were satisfied with the results of the program and will continue to use and perhaps expand P4P (Rosenthal et al., 2007).

Pay-for-performance (P4P) is a much touted tool presently in use around the world. The literature is replete with publications which both support and refute its utility (Kautter et al., 2008; Lindenauer et al., 2007; Meterko et al., 2006; Rosenthal and Dudley, 2007). Models for P4P programs exist across commercial health plans and Centers for Medicare and Medicaid Services and vary on breadth and depth of performance measurement. Issues related to measure selection, measure reliability, the financial impact on providers and therefore the ability to incent behavior have been well outlined and challenge P4P as a useful intervention (Rosenthal et al., 2007). Reliability has surfaced in the literature as one indication of the ability of a measure to differentiate physicians and has raised the issue of misclassification and robustness of the P4P models to detect real differences in performance among physicians rather than any shortcomings of the measurement (Hofer et al., 1999; Scholle et al., 2008). Health care organizations and physician groups are left to manage the sometimes conflicting reports across health plans, cynical physicians, and contract negotiations on non-standardized measures and non-transparent weighting schemes (Draper, 2009). While some health care organizations have been developing their own performance tools like Balanced Scorecards (Impagliazzo et al., 2009; Curtright et al., 2000; Stewart and Greisler, 2002), a standardized methodology for evaluating physician

performance measures across domains of care, practice functionality and patient experience is lacking.

Physician profiling is a mechanism for holding physicians accountable for the care delivered to their patient panel. Current programs typically select a set of measures and then either combine the measures or use individual measures to determine payout incentives or to use for public reporting. The performance feedback provides an opportunity for physicians, clinics and organizations to study how they deliver care, rather than just results of an individual episode, and to benchmark their performance relative to their peers. Garrett (2007) found that if physicians discover or are provided with credible, actionable information that points to an area in their practice where value provided to their patients can be improved, they will want to do so. The question then becomes whether the feedback is useful or valuable to physicians and other users of the profiling program.

2.2.2 Criteria for Physician Profiling Models

The value of physician performance ranking depends on several factors: (1) the validity of the measure in describing performance of the ranking physicians; (2) the selection of the measures to satisfy the purpose of the analysis; (3) the completeness and accuracy of the data; (4) the reliability of the profiles; and (5) the usefulness to the end users. These factors are essential in building an assessment tool to ensure fair comparisons across physicians.

2.2.2.1 Validity

The validity of the measure for profiling questions whether the chosen measure is a true indicator of the performance of the individual physician or whatever the profiling program is trying to achieve. This is an important property of physician profiling because validity can sometimes suggest action ability. There might be unwanted influences on the results not

controlled for by the physician, such as clinic functionalities and patient panel characteristics, shown to be significant predictors of quality and efficiency (Hofer et al, 1999). For instance, the presence of registries and reminder systems for chronic and preventive care might have more influence on chronic and preventive measures than that of an individual physician. Solberg et al. (2008) found a significant correlation between the presence of practice systems (e.g. decision support and delivery system redesign) and process and outcome measures of diabetes mellitus. Focusing on practice organization functionalities might be more valuable in understanding opportunities to improve quality of care than individual and in some cases disease specific measures, hence a new focus on Medical Home (Rogers, 2008; Rosenthal, 2008; Vesely, 2008). A study of 12,110 patients with diabetes from 258 PCPs, 42 provider groups, and 13 VA facilities found the greatest amount of variance tended to be attributable to the facility level (Krein et al., 2002).

Case-mix adjustments factor out the variability in the data caused by known influences such as patient demographics, socioeconomic status, and duration of disease, which may be desirable when variation exists in patient characteristics across the individual physicians. Case-mix adjustment can provide a more vivid picture of the variation in the measure caused by physicians alone. Hofer et al. (1999) found that no more than 4% of total variation was caused by individual physicians, after case mix adjustment, which raises the question of whether the effort involved in profiling is worth the benefit. This study does not seek to solve these validity issues, but to raise the concern that measures should be chosen based on how well they measure whatever the profiling program is trying to achieve.

2.2.2.2 Purpose

Selecting measures to satisfy the purpose of the analysis is another dimension of value. Using a broad set of measures might provide a better understanding of overall quality of care delivered across a physician's entire patient panel versus a select disease-specific subgroup of patients (Hofer et al.,1999; Parkerton et al., 2003). With a narrow set of measures, some primary care specialties like pediatrics might be under-represented in the P4P calculations creating little or no incentive for improvement by individual physicians (Scholle et al., 2009). The source of the data also influences what is being measured. Relying on claims data hinders the types of clinical actions that can be profiled, focusing largely on process measures rather than on intermediate clinical outcomes. For example, a profiling program focusing on intermediate clinical outcomes might consider a common test for monitoring long term blood sugar control in patients with diabetes mellitus (hemoglobin HbA1c). Claims data would only indicate the test was performed but not whether the patient is clinically in control. Rosenthal et al. (2007) studied twenty-seven health plans with P4P programs and found they emphasized efficiency and outcome measures rather than process measures. Additionally, current profiling tools sometimes use non-standardized measures and weighting schemes that are non-transparent to individual physicians, which makes it difficult for individual physicians to make appropriate decisions about changing the way they work (Draper, 2009). Our model in this study was designed to be flexible in using measures across multiple sources of data and domains of care to minimize the influence of any one measure, to cover physicians across specialties, and to provide a breadth of care functionalities.

2.2.2.3 Completeness and accuracy of the data

The completeness and accuracy of the data permits a comprehensive assessment of an individual physician. Most health plans have access to only their claims, which may not be a representative sample of a physician's patient population and yield insufficient sample sizes to support reliable profiling. As a result, some physicians receive conflicting ranking reports from different health plans, resulting in inconsistent feedback (Draper, 2009).   Scholle et al. (2009) suggested pooling physician data across payers with the ability to link patients to physicians across payers. There are independent coalition groups like The Massachusetts Health Quality Partners (MHQP) which address this issue by collecting results from multiple health plans in a market and publically disseminate performance scores against benchmarks (http://www.mhqp.org/, accessed on October 14, 2009). Our model will build on this idea of pooling patient level data across payers for each individual physician to assist in the development of a reliable metric.

Data errors can also reduce the accuracy of physician profiling. For example, an analysis of medical records versus Medicaid paid claims data found a 2.6% discrepancy in total volume of visits (Steinwachs et al., 1998).

2.2.2.4 Reliability

Reliability measures the ability to confidently discriminate one individual physician from another and is influenced by the differences between physicians, sample size and measurement error. High reliability does not indicate good physician performance but rather the ability to differentiate one physician's performance relative to that of other physicians. If the reliability is low, one cannot distinguish between the performance of physicians for that measure. Reliability can be low even when there is very little physician variability in a compressed scale (even if data

is accurate and complete), when there are too few physicians to distinguish, or when there are small sample sizes or quality events. Reliability also varies by the rate of the measure as dictated by the binomial equation for measurement error. Reliability is the ratio of the physician-to-physician variance to the sum of the physician-to-physician variance and the average within physician variance. A reliability estimate greater than 0.70 has been typically deemed as an acceptable threshold by psychometric theory to differentiate physicians (Nunnelly and Bernstein, 1994). Hence, sample size can be determined to obtain desired reliability estimate, given an average measure rate and known physician-to-physician variation. Krein et al (2002) found that after case-mix adjustment, small physician variability requires larger sample sizes to yield a desired reliability estimate. Scholle et al. (2008) found that most quality measures require at least fifty quality events to yield an acceptable reliability estimate and proceeded to introduce the idea of composite measures as a mechanism to address low reliability. Nyweide et al. (2009) found that very few practices had enough Medicare patients to reliability detect a 10% difference on common individual measures against a national benchmark and recommends novel approaches for amassing quality events. Our proposed model will address the issue of low reliability by pooling data across multiple payers and data sources and combining measures into domains of care.

2.2.2.5 Usefulness

Usefulness is in part an outcome of the above criteria but also depends on who is defined as the end user. The primary end users of physician-oriented P4P programs are health plans, who use the results to determine financial consequences of performance. While physician groups have some say on measure selection and the performance targets during contract negotiations, they often are unaware of specific model variables, including case-mix adjustments, weighting

schemes, sample size thresholds or reliability considerations. Improving quality of care with

P4P is only attainable if physicians value the feedback as a fair assessment of their performance.

Our model methodology encompasses four attributes suggested for the usefulness of physician

profiling: flexibility, end-user involvement, transparency of the details and fairness to targeted

physician groups (Garnick et al., 1994).

2.3 Methods

We use the process shown in Figure 2.1 to build our model. The process consists of five

steps: modeling specification, model fitting, model assessment, model validation, and plan for

output. In this section, we describe each step of the model building process; in the next section,

we present its application to a case study site (Figure 2.2) and the results.
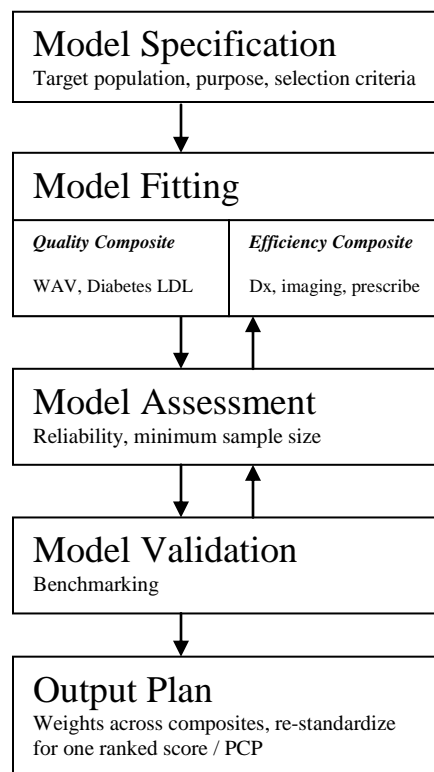
Figure 2.1 Process of Model Building

2.3.1 Model Specification

Model specification is essential for achieving ***usefulness***, and involves establishing the target population, defining the ***purpose*** of the profiling analysis, and identifying selection criteria for measure inclusion and ***validity***. Selection criteria may include but not limited to case-mix adjustable, feasibility of data access, use in P4P programs, use in public reporting, evidence based, and clinically relevant. An evaluation of potential measures against the required selection criteria is recommended.

2.3.2 Model Fitting

Model fitting is the process of building the model by deciding which measures to add, across what domains and how to combine domains. This step is critical for meeting the model criteria of ***validity***, ***completeness and accuracy of data*** and ***usefulness***. A list of potential measures might include P4P metrics, Medical Home survey results, patient experience survey results, internal reward program metrics, and Healthcare Effectiveness Data and Information Set (HEDIS). HEDIS contains 71 physician measures for different primary care specialties across eight domains of care including effectiveness of care, availability of care, satisfaction with the experience of care, use of services, cost of care, and health plan measures (NCQA, 2007). These standardized measures were designed for the comparison of health plans, and can be ***useful*** for comparing physicians, practices, and groups.

Selected measures should be sorted into homogeneous domains (e.g., chronic, acute, preventive, resource utilization, Medical Home elements and patient experience) to satisfy the purpose of the analysis. Additional measures should be considered to diversify measures across PCP specialties, across domains of care and across data sources (e.g., lab data, EHR and patient surveys) in order to ensure fair and accurate comparisons across physicians and to ensure meaningful use of the results. However, the ***completeness and accuracy of the data*** for each

measure can minimize the extent of which measures can be added along with the *validity* of the measure to appropriately measure physician performance in the domain.

The measures in each domain are then combined quantitatively to form composite measures. Composite measures allow more physicians to be reliably assessed than individual measures due to an increase in sample size (Scholle et al., 2008; Caldis, 2007). Several computational techniques have been used to create composites. One simple technique is to combine all of the numerators and divide by all of the denominators for each physician and in each domain. A second computational method creates an aggregate measure by taking the mean of the component measures (Parkerton et al., 2003). A third method introduced by Scholle et al. (2008) creates a composite measure using weighted Z scores across measures per individual physician. For each domain, the resulting composite score is then standardized across physicians to yield a final ranking. A detailed description of the model is provided in Chapter 3.

2.3.3 Model Assessment

Model assessment is the process of studying the composites for *reliability*, determining minimum sample sizes and if necessary looping back to model fitting to determine optimal domain structure to meet a minimum reliability of 0.70. This study uses a reliability methodology suggested by Adams (2009) and applied by Scholle et al. (2008) and Adams (2010). First, the physician-to-physician variance is determined using hierarchical linear modeling for each composite using a PROC MIXED procedure in SAS (Adams, 2009). The composite error variance was calculated as the summation of the squared weights multiplied by the individual measure average error variance. An error variance was calculated per physician and is based on the quality event distribution and the rates of the individual measures for that provider. Reliability is calculated by dividing the physician-to-physician variance by the total of

the physician-to-physician variance and the within physician variance using the Spearman-Brown formula (Shrout and Fleiss, 1979). The assumption is as sample size of a composite measure increases, the reliability approaches one. The reliabilities for each physician are calculated and plotted in order to determine the minimum size. If the minimum reliability is less than 0.70, it is necessary to loop back to the Model Fitting step and either add additional measures to increase the total sample size per physician or to reassess the domain structure. The reliability calculations are detailed in Chapter 3.

2.3.4 Model Validation

Model validation involves studying individual and composite measure results against known benchmarks and if necessary, dropping measures when causes of discrepancies cannot be determined. The intent is to screen for large discrepancies or biases due to systematic errors in data collection or measurement. This step is critical for assessing the ***completeness and accuracy of the data***. Quality hits or misses can be double-checked by conducting a chart review on a randomized sampling of individual patients and select measures. New calendar year data can also be compared against previous years with the caution that changes can occur in standardized measurement rules like HEDIS. Measurement results can be benchmarked against externally vetted data particularly that collected by health care coalitions that pool data across payers. Discrepancies in measure results between the benchmark and the newly calculated measures are inevitable due to timing, data sources, and physician-patient assignment. Rules for percent deviation from benchmarks should be established and action taken to troubleshoot deviations above the threshold. Measures should be dropped when root causes cannot be determined, in which case model fitting and/or model assessment steps are revisited.

2.3.5 Output Plan

The last stage of physician profiling is developing an Output Plan, in which the organization assigns weights to each domain in alignment with their organizational policy; hence strengthening the *usefulness* of the model. The final weighted scores across domains are combined and then standardized to create a final ranking for all physicians across the organization. Payout thresholds can be determined based on final scores. Payout can then be determined as a function of individual physician score, panel size and allotted dollars.

2.3.6 Discussion

We use a process to build our model designed to support the value criteria suggested in the literature. The value of the physician profiling model depends on how factors related to validity, purpose, completeness and accuracy of data, reliability and usefulness are interweaved and engrained in the building process. For example, the usefulness of the model is not only dependent on the success of the other factors but also on the recognition and involvement of the end users as described during the Model Specification and Output Plan steps. These factors guide the model building process and also frame the limitations of and future studies for physician profiling.

2.4 Application To Case Study And Results

2.4.1 Case Study Setting and Data

The performance profiling model presented in this study was used to profile primary care physicians in a managed care network. The study population included 199 primary care specialists across pediatric primary care, family medicine and general internal medicine from 68 Massachusetts outpatient clinics. All physicians had provided ambulatory care for at least 12 months in 2008.

The case study site, Health Clinic, provided the 2008 measure rates, numerators and denominators per individual physicians for selected measures. Data was pulled from a central repository consisting of claims data from three health plans, lab results, and billing/scheduling data. Health plan administrative claims and enrollment data across three payers represented roughly 75% of all paid claims. Additional information included panel size per physician, physician specialty, employment status per physician and externally vetted benchmark performance per measure.

2.4.2 Resulting Model

The final model design for Health Clinic is illustrated in Figure 2.2. The target population was all physicians who contributed to the network rate for 2008 performance. A sub-population of primary care physicians across primary care pediatrics, general internal medicine and family medicine was eligible for an incentive reward payout. Eligibility criteria required that the physician be in the network for the full 12 months in 2008 and active at the time of distribution, with a panel size within the analytic file greater than 30. The purpose of profiling was to reward physicians across domains of quality and efficiency with a payout based on their contribution to the network. A total of 265 providers including infectious disease and nursing home specialties were included in the individual measure analyses of which only 199 primary care physicians (110 internal medicine, 56 family medicine and 33 pediatric) met the criteria for the target population and payout. Including all providers enabled comparison between the organizational rate and other benchmarks.
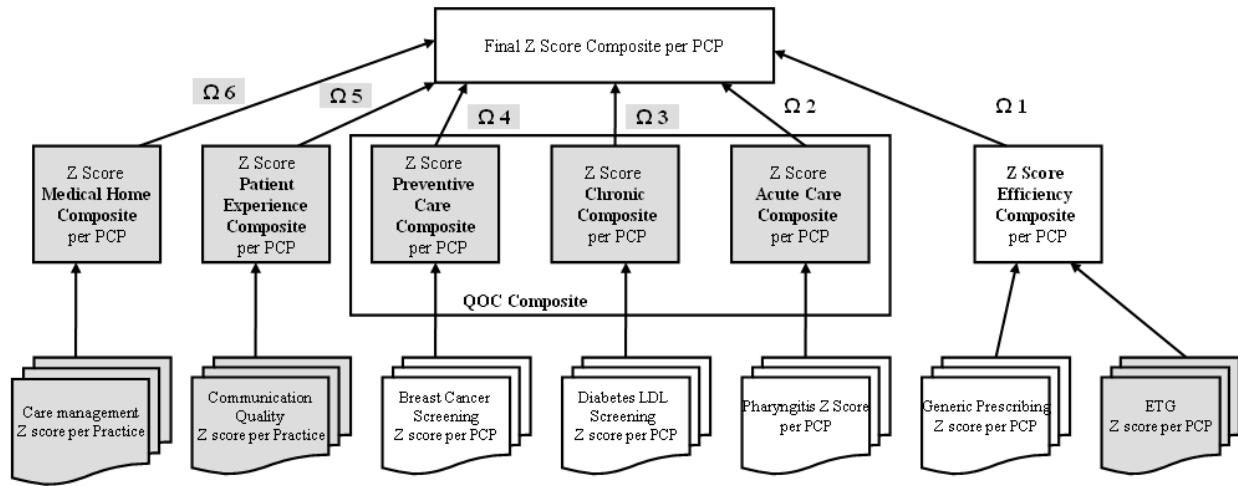
Figure 2.2 Model Design at Case Study Site

The selection criteria for measure inclusion consisted of: feasibility of data access; P4P measurement; publically reported; externally vetted; evidence based; clinically relevant; and available benchmarks. Forty-two quality and two efficiency measures were evaluated against the selection criteria. Twelve measures were dropped due to small physician coverage and six measures were dropped due to nonstandard measurement. Twenty-six remaining measures met the initial criteria and were grouped into domains of quality of care and efficiency. In computing a composite score per physician, measures were only used if they had more than four quality events in order to minimize variability and multimodal distributions. With this threshold, three measures were dropped due to an inadequate number of physicians to profile.

The individual measure results were validated against 2007 Massachusetts Health Quality Partners (MHQP) network results and EHR data. Two measures, eye screening for diabetic retinal disease and well infant visits, were dropped due to discrepancies greater than 10% that could not be resolved. Model validation was also carried out with each physician specialty and physician clinic type to search for major biases indicating inaccurate or incomplete data. The

final list of 21 measures and 2 domains is shown in Table 2.1, along with the number of physicians having quality events for that measure. The number of physicians varied within primary care specialties. Diabetes mellitus measures had more physicians represented and a higher mean number of quality events than coronary artery disease (CAD) measures. Cancer Screening measures had the highest mean number of quality events across physicians.

The domains were then added using weights determined by the governing committee at Health Clinic. The final weighted scores across domains were standardized to create a final overall composite ranking score for all physicians across the organization as seen in Figure 2.2.

The model was implemented for the 2008 Rewards program at Health Clinic. Payouts were distributed based on the overall composite ranking score per physician and their panel size. The payout plan excluded all primary care physicians who had a final Z score less than -1. Those physicians who did not meet the original eligibility requirements were also dropped from this list at this point. The overall composite ranking score for each primary care physician was added to one and then multiplied by their panel size to yield panel points per physician. The total available incentive dollars was then divided by the total panel points across all physicians to yield a $/panel point rate. This rate was then multiplied by the panel size of each physician to yield their individual payout quantity.

2.4.3 Model Evaluation

Table 2.1 summarizes the reliability results for the individual measures calculated at the mean rate and mean quality events across providers. The individual reliabilities ranged from 0.11 to 0.89 with generic prescribing having the highest reliability, highest average number of quality events, and highest percent of physicians meeting the minimum sample size requirement for reliability of 0.70. Colorectal cancer screening and Well Adolescent Visit measures also had

reliability estimates above 0.70. Overall, the reliability results varied considerably across the individual measures and typically improved with an increase in quality events. Accordingly, the sample size needed to achieve a reliability of 0.70 at the mean rate of performance also varied by measure and ranged between 11 for Pharyngitis and 233 for Well Child Visit. As expected, as the reliability of the measures increased, the percent of physicians meeting the minimum sample size increased.

Table 2.1 Table of Reliability Results for Individual Measures in Each Domain

Domain	Measure	Total Number of Physicians with Quality Events >4	Mean No. of Quality Events across Physicians	Reliability at Mean Quality Events and Mean Rate	Sample Size Needed at mean rate for 0.70 Reliability	% of physicians meeting minimum sample size for 0.70 reliability
Quality of Care	Well Adolescent Visit	207	34	0.73	29	38
	Well Child Visit	97	24	0.19	233	0
	Pharyngitis	82	7	0.62	11	0
	Upper Respiratory Infection	104	11	0.33	51	6
	Chlamydia Ages 16-20	142	6	0.23	52	0
	Chlamydia Ages 21-24	138	6	0.20	58	0
	Diuretics	172	16	0.60	26	25
	ACE and ARBs	179	21	0.68	24	44
	CAD LDL Control	104	5	0.11	98	0
	CAD LDL Testing	100	7	0.16	78	0
	Diabetes Nephropathy	181	14	0.41	46	1
	Diabetes LDL Control	163	10	0.22	82	0
	Diabetes LDL Testing	181	14	0.26	90	0
	Diabetes A1C < 7 (Good)	161	11	0.39	39	3
	Diabetes A1C >9 (Poor)	161	11	0.15	148	0
	Diabetes Testing (2/yr)	181	14	0.37	54	0
	Cervical CS	172	68	0.53	138	22
	Breast CS	173	60	0.54	117	20
	Colorectal CS	154	70	0.78	45	59
Efficiency	Lower Back Pain	166	6	0.17	70	0
	Generic Prescribing	259	1111	0.97	77	89
Minimum across measures		82	5	0.11	11	0
Maximum across measures		259	1111	0.97	233	89

The composite reliabilities were estimated for each physician in the domains of quality of care and efficiency. The results are summarized in Table 2.2; all 265 physicians were included in both composites for profiling. The mean number of quality events in the quality of care composite exceeded all of the means of its components. The median reliability for the quality of care composite was 0.98 and the median reliability for the efficiency composite was 0.97, both

exceeding the recommended threshold of 0.70. The approximate sample size needed to achieve

a reliability estimate greater than 0.70 was 17 for the quality of care composite and 63 for the

efficiency composite compared to the individual measures that ranged between 11 and 233.

Table 2.2 Reliability Results for the Composites

Composites	Total Physicians	Mean Quality Events Across Physicians	Average Physician Rate	Median Reliability across physicians	Approximate Sample Size Needed for 0.70 Reliability	% of physicians meeting minimum sample size for 0.70 reliability
Quality of Care	265	252	0.72	0.98	17	96
Efficiency	265	1115	0.77	0.97	63	90

The percent of total evaluated physicians (265) meeting the minimum reliability threshold

was 96% for the quality of care and 90% for the efficiency; of which, 100% of all physicians

receiving a payout had reliability estimates greater than 0.70.

2.5 Discussion

2.5.1 Study Implications

The physician profiling model introduced in this paper is expandable to include additional

data sources beyond administrative claim data and additional domains for patient experience and

practice functionalities, making it flexible for diverse health care settings. Health Clinic has

future plans to expand domains in their model to include internally assessed Medical Home

survey data, patient experience survey data administered by MHQP, Episode Treatment Group

data and an expansion of the quality of care domain into three domains of care, acute, preventive

and chronic (Figure 2). The hurdle to expanding the quality of care domain is finding additional

measures that meet the model criteria of *validity*, *completeness and accuracy of data* and

reliability to confidently assess physicians.

Once the well known, heavily populated measures were consumed, the smaller less populated measures became the target of inclusion, possibly introducing additional variability and multimodality in the data.   The model design steps of Model Fitting, Model Assessment and Model Validation were iterated to add less common or disease or treatment specific measures. Although additional measures are needed to increase quality events in the composites for *reliability*, adding multimodal data across a small group of physicians can be problematic. We found that the Z score range across physicians per measure was highly influenced by the minimum quality event threshold and the number of physicians profiled. By definition of a Z score (Petruccelli, 2007), as the variation in a measure increases, the Z score range becomes compressed. A discussion of minimum thresholds for quality events and number of physicians is explored in Chapter 3.

Out of the original 44 measures identified for possible inclusion in the model, 15 measures were eventually eliminated due to limited access to data causing poor physician coverage and small denominator sizes. While patient panel size and mix per physician are major contributors, additional payer data from sources like Medicare and Medicaid data would support a higher number of quality events per physicians and the ***completeness and accuracy of data***. Additional data sources like Electronic Medical Records (EMR) also have the potential to provide additional data that would support expansion of measures in the quality of care composite. Psychometric instruments offer the ability to measure softer outcomes like health status, function, and comfort (Berwick, 2009). While the use of patient data from electronic health records appears promising, Linder et al. (2009) found that accurate identification of pneumonia visits using an EHR was challenging.

Benchmarking the results to an external source proved valuable because it highlighted measurement errors and possible discrepancies in interpreting HEDIS guidelines but was also challenging. For instance, MHQP used correction factors to accommodate measurement and sampling error, thereby adding an intentional bias in the results. There may have also been differences in the number of payers used and the ability to track patients across physicians and health plans, affecting the determination of continuous enrollment. And finally, if comparing against historical data, the possibility that individual measures improved exists along with possible modifications to the inclusion and exclusion codes for the HEDIS measure.

The model was adopted at Health Clinic as a payout methodology for the second annual internal incentive reward program and remains in *use*. The previous reward program ranked physicians on a small set of P4P measures and organizational incentives and lacked consistency across the members of the network. . Additionally, the program did not measure all of the physicians in the managed care network. A sensitivity analysis on the newly adopted model was conducted to understand the impact of undetected errors in the quality hits per measure on payout. If twenty quality hits were missed for the highest weighted measure, an approximate 2% change in payout would occur for those physicians with network panel sizes in the top quartile and a 3% change would occur for physicians having an average network panel size. The payout methodology also minimized the risk of misclassification by ranking physicians across a continuous function versus grouping physicians into levels for payout.

2.5.2 Limitations

The model ranks physicians based on annual measures that rely heavily on claims data that has a typical run out period of six months. Therefore, it can be as long as nine months after the performance period ends that the physicians actually get their report card. The delay in feedback

can hinder agility to respond to performance issues and therefore, the ***usefulness*** of the model. Consequently, performance improvement from interventions might take two years to actually show in the results.

Another limitation of the model in its current state is its inability to recognize performance improvement over previous years and its inability to set benchmarks for future performance. The current model simply ranks physicians based on their performance in one cycle. Tompkins (2009) develops a Value-Based Purchasing model which calculates a total performance score by using selected benchmarks and minimum attainment levels per measure. .

This model does not allow for action at the individual measure level but rather at the domain level. While the physicians were given their quality hits and quality events per measure, the results cannot be interpreted to be statistically significant. Also, any controversy regarding ***validity***, clinical relevance and evidence based assessment of individual measures are not the focus of this paper. Therefore, the model does not address the shortcomings of individual measures themselves but rather offers a construct to combine them within domains of care or functionality, thereby, minimizing the influence of any one measure. Having a birds-eye view of the delivery of care across domains might address major deficiencies in the health care processes versus viewing care through a keyhole of select measures.

The ***purpose*** of the physician profiling was to reward physicians for their contribution to the network rates. However, due to variation in patient panel sizes, a physician who performed above the network rate could receive a standardized score below the physician average. The panel size contribution is factored in the payout calculation after the ranking has been established. This allowed for physicians with smaller panel sizes (>30) to be recognized for their performance in caring for their patients regardless of the size of their patient panel.

As implemented at Health Clinic, the model design did not attempt to rank physicians with a panel size less than 30 or include measures with less than 30 physicians. However, alternative approaches for computing composite scores by combining similar measures (adding numerators and denominators) was suggested and should be considered when measures share like functionalities.

Creating a model to ensure fair and accurate comparisons across physicians and to ensure appropriate use is a challenge given the availability of data, limiting the selection of measures and the ***reliability*** of profiling tool. In particular, the case study site did not have access to claims data from additional payers (e.g., Medicaid and Medicare, PPO products of three included payers, and other payers). Initiatives to create data transparency across payers at the individual physician performance level can improve the ***completeness and accuracy of data*** across patient populations and potentially improve the ability to profile physicians, clearly identify domains for improvement and ultimately improve the quality and efficiency of health care delivery.

2.5.3 Future Research

The model design as implemented at Health Clinic did not use case-mix adjustments or stratify for known variants like patient demographics, socioeconomic status or medical practice characteristics. Biases due to patient compliance and practice functionalities were assumed constant across all sites. Future studies should explore expanding the model design to include these considerations, particularly when the target population is diverse in location, practice characteristics and patient types. While medical practice characteristics could play a significant role in the quality of care delivered to patients and the efficiency of services, Landon et al. (2008) showed that only financial incentives, EMRs and care management approaches had an effect on quality of care measures for coronary artery disease (CAD) patients.

The model is flexible enough to investigate opportunities to include a hierarchical weighting process where clinics and physicians are part of a goal setting process and choose their own focus for improvement. However, those weights would have to be determined well in advance of knowing the previous cycle results or "gaming" the system could take place.

2.6 Conclusions

In this study, we presented a methodology for creating a physician profiling model designed to support key criteria suggested in the literature as valuable to physician profiling. We then developed a physician profiling model at Health Clinic with an emphasis on the ability to reliably differentiate the performance of one physician from another. To date, it has been difficult to develop reliable profiling tools at the individual practitioner level due to small denominators which has resulted in very narrow sets of measures. We have devised and tested a model which allows a broad set of measures to be used at the individual or group level. Using P4P to stimulate an improvement in quality of care by holding physicians accountable for the care provided to their patients is only attainable if physicians value the feedback as a fair assessment of their performance. There is some evidence that making the right information available and changing payment incentives does support improved care (Lexa, 2008).

By designing the value criteria into the model building process, we have also provided other researchers with a critical lens to evaluate physician performance models and identify the limitations for future studies. While the resulting model at Health Clinic met the design expectations of reliability, purpose, and usefulness, more work is necessary to expand the model across additional measures and domains, especially for validity and completeness and accuracy of data. In particular, health care initiatives to create shared pools of transparent physician performance information for Medicare, Medicaid, and all private insurers would advance the

completeness and accuracy of data and therefore, the ability to assess and profile physician performance. And, secondly, the creation of valid measures of performance across medical conditions would broaden the eligible patient population per measure and hence support reliable profiling.

We presented a model that was built with the input and approval of the governing body of the profiled physicians. The model is flexible and can accommodate change in future years, is transparent in terms of measurement, aligns with the organizational quality agenda and is perceived as fair to targeted physician groups, thereby supporting the key elements needed for usefulness of physician profiling (Garnick et al., 1994). We developed a hierarchical composite model which both broadens the array of measures as well as facilitates the assignation of dollars to more robustly incent and reward physician performance.

2.7 References

Adams, J. L. "The Reliability of Provider Profiling: A Tutorial," *Rand* 2009.

Adams, J. L., A. Mehrotra, J. W. Thomas, and E. A. McGlynn. "Physician Cost Profiling--Reliability and Risk of Misclassification," *The New England Journal of Medicine* (362:11), Mar 18, 2010, pp. 1014-1021.

Berwick, D. M., and S. King. "Web Institute of Health Improvement: Road to a Better Health Care System," September 24, 2009. Available from: http://www.ihi.org/IHI/Programs/AudioAndWebPrograms/WIHI.htm.

Berwick, D. M. "Measuring Physicians' Quality and Performance: Adrift on Lake Wobegon," *JAMA: The Journal of the American Medical Association* (302:22), December 9, 2009, pp. 2485-2486.

Caldis, T. "Composite Health Plan Quality Scales." *Health Care Financing Review* (28:3), 2007, pp. 95.

Centers for Medicare and Medicaid Services (CMS). "National Health Expenditures: 2007 Highlights," 2009.

Christianson, J. B., P. B. Ginsburg, and D. A. Draper. "The Transition from Managed Care to Consumerism: A Community-Level Status Report," *Health Affairs (Project Hope)* (27:5), Sep-Oct, 2008, pp. 1362-1370.

Curtright, J. W., S. C. Stolp-Smith, and E. S. Edell. "Strategic Performance Management: Development of a Performance Measurement System at the Mayo Clinic," *Journal of Healthcare Management / American College of Healthcare Executives* (45:1), Jan-Feb, 2000, pp. 58-68.

Draper, D. A. "Physician Performance Measurement: A Key to Higher Quality and Lower Cost Growth Or a Lost Opportunity?" *Health Systems Change Brief* (3),June, 2009.

Garnick, D. W., J. Fowles, A. G. Lawthers, J. P. Weiner, S. T. Parente, and R. H. Palmer. "Focus on Quality: Profiling Physicians' Practice Patterns," *The Journal of Ambulatory Care Management* (17:3), Jul, 1994, pp. 44-75.

Garrett, K. "The Measurement of Health Care Performance: A Primer from the CMSS," 2007.

Hofer, T. P., R. A. Hayward, S. Greenfield, E. H. Wagner, S. H. Kaplan, and W. G. Manning. "The Unreliability of Individual Physician "Report Cards" for Assessing the Costs and Quality of Care of a Chronic Disease," *JAMA: The Journal of the American Medical Association* (281:22), June 9, 1999, pp. 2098-2105.

Impagliazzo, C., A. Ippolito, and P. Zoccoli. "The Balanced Scorecard as a Strategic Management Tool: Its Application in the Regional Public Health System in Campania," *The Health Care Manager* (28:1), Jan-Mar, 2009, pp. 44-54.

Institute of Medicine (U.S.). Committee on Quality of Health Care in America. *Crossing the Quality Chasm : A New Health System for the 21st Century*, National Academy Press, Washington, D.C., 2001.

Kaiser Family Foundation. "Side-by-Side Comparison of Major Health Care Reform Proposals," *Focus on Health Reform,* September 18, 2009.

Kautter, J., G. C. Pope, M. Trisolini, and S. Grund. "Medicare Physician Group Practice Demonstration Design: Quality and Efficiency Pay-for-Performance," *Health Care Financing Review* (29:1), Fall, 2008, pp. 15-29.

Landon, B. E., S. L. Normand, E. Meara, Z. Qi, S. R. Simon, R. Frank, and B. J. McNeil. "The Relationship between Medical Practice Characteristics and Quality of Care for Cardiovascular Disease," *Medical Care Research and Review : MCRR* (65:2), Apr, 2008, pp. 167-186.

Lexa, F. J. "Pay for Performance and the Revolution in American Medical Culture," *Journal of the American College of Radiology : JACR* (5:3), Mar, 2008, pp. 168-173.

Lindenauer, P. K., D. Remus, S. Roman, M. B. Rothberg, E. M. Benjamin, A. Ma, and D. W. Bratzler. "Public Reporting and Pay for Performance in Hospital Quality Improvement," *The New England Journal of Medicine* (356:5), February 1, 2007, February 1, 2007, pp. 486-496.

Linder, J. A., E. O. Kaleba, and K. S. Kmetik. "Using Electronic Health Records to Measure Physician Performance for Acute Conditions in Primary Care: Empirical Evaluation of the Community-Acquired Pneumonia Clinical Quality Measure Set." *Medical Care* (47:2), 2009, pp. 208-216.

Massachusetts Health Quality Partners (MHQP). "Quality Insights: Healthcare Performance in Massachusetts," Available at: http://www.mhqp.org/. Accessed October, 14, 2009.

Meterko, M., G. J. Young, B. White, B. G. Bokhour, J. F. Burgess, D. Berlowitz, M. R. Guldin, and M. Nealon Seibert. "Provider Attitudes Toward Pay-for-Performance Programs: Development and Validation of a Measurement Instrument." *Health Services Research* (41:5), 2006, pp. 1959.

National Committee for Quality Assurance. "HEDIS 2007 Technical Specifications for Physician Measurement," Washington, DC: National Committee for Quality Assurance: 2007.

National Committee for Quality Assurance. "Physician Practice Connections—Patient-Centered Medical Home (PPC-PCMH) Standards," Available at: http://www.cms.hhs.gov/DemoProjectsEvalRpts/. Accessed May, 28, 2009.

Nyweide, D. J., W. B. Weeks, D. J. Gottlieb, L. P. Casalino, and E. S. Fisher. "Relationship of Primary Care Physicians' Patient Caseload with Measurement of Quality and Cost Performance," *JAMA : The Journal of the American Medical Association* (302:22), Dec 9, 2009, pp. 2444-2450.

Parkerton PH, Smith DG, Belin TR, and Feldbau GA. "Physician Performance Assessment: Nonequivalence of Primary Care Measures." *Medical Care* (41:9), Sep-1, 2003, pp. 1034-1047.

Petruccelli, J. D., B. Nandram, and M. Chen. *Applied Statistics for Engineers and Scientists*, Prentice Hall, Upper Saddle River, N.J., 1999.

Rogers, J. C. "The Patient-Centered Medical Home Movement--Promise and Peril for Family Medicine." *Journal of the American Board of Family Medicine (Formerly Journal of the American Board of Family Practice)* (21:5), 2008, pp. 370.

Rosenthal, M. B., and R. A. Dudley. "Pay-for-Performance: Will the Latest Payment Trend Improve Care?" *JAMA : The Journal of the American Medical Association* (297:7), Feb 21, 2007, pp. 740-744.

Rosenthal, M. B., F. S. de Brantes, A. D. Sinaiko, M. Frankel, R. D. Robbins, and S. Young. "Bridges to Excellence--Recognizing High-Quality Care: Analysis of Physician Quality and Resource use." *Am J Manag Care* (14:10), 2008, pp. 670.

Rosenthal, T. C. "The Medical Home: Growing Evidence to Support a New Approach to Primary Care." *Journal of the American Board of Family Medicine (Formerly Journal of the American Board of Family Practice)* (21:5), 2008, pp. 427.

Scholle SH, Pawlson LG, Solberg LI, Shih SC, Asche SE, Chou AF, and Thoele MJ. "Measuring Practice Systems for Chronic Illness Care: Accuracy of Self-Reports from Clinical Personnel." *Joint Commission Journal on Quality and Patient Safety* (34:7), Jul-1, 2008, pp. 407-416.

Scholle, S. H., J. Roski, J. L. Adams, D. L. Dunn, E. A. Kerr, D. P. Dugan, and R. E. Jensen. "Benchmarking Physician Performance: Reliability of Individual and Composite Measures," *The American Journal of Managed Care* (14:12), Dec, 2008, pp. 833-838.

Scholle, S. H., J. Roski, D. L. Dunn, J. L. Adams, D. P. Dugan, L. G. Pawlson, and E. A. Kerr. "Availability of Data for Measuring Physician Quality Performance," *The American Journal of Managed Care* (15:1), Jan, 2009, pp. 67-72.

Shrout, P. E., and J. L. Fleiss. "Intraclass Correlations: Uses in Assessing Rater Reliability," *Psychological Bulletin* (86:2), 1979, pp. 420-428.

Solberg, L. I., S. E. Asche, L. G. Pawlson, S. H. Scholle, and S. C. Shih. "Practice Systems are Associated with High-Quality Care for Diabetes," *The American Journal of Managed Care* (14:2), Feb, 2008, pp. 85-92.

Steinwachs, D. M., M. E. Stuart, S. Scholle, B. Starfield, M. H. Fox, and J. P. Weiner. "A Comparison of Ambulatory Medicaid Claims to Medical Records: A Reliabiligy Assessment," *American Journal of Medical Quality* (13:2), June 1, 1998, pp. 63-69.

Stewart, L. J., and D. Greisler. "Measuring Primary Care Practice Performance within an Integrated Delivery System: A Case Study," *Journal of Healthcare Management* (47:4), Jul, 2002, pp. 250.

Tompkins, C. P., A. R. Higgins, and G. A. Ritter. "Measuring Outcomes and Efficiency in Medicare Value-Based Purchasing," *Health Affairs (Project Hope)* (28:2), Mar-Apr, 2009, pp. 251-261.

Vesely, R. "Medical Home Push. Align Payment with Quality: AHIP." *Modern Healthcare* (38:26), 2008, pp. 15.

Wodchis, W. P., J. S. Ross, and A. S. Detsky. "Is P4P really FFS?" *JAMA : The Journal of the American Medical Association* (298:15), Oct 17, 2007, pp. 1797-1799.

3.0 Physician Profiling: Detailed Calculations and Future Study Discussion

3.1 Introduction

Physician profiling is gaining attention as legislators, health care organizations, insurance companies and payers strive to improve quality, safety, and efficiency of our health care system. In chapter 2, we developed a physician profiling model that addresses the reliability shortcomings of current profiling tools. Reliability has been identified as a characteristic of a measure that gauges the ability to differentiate physicians.   This chapter describes the detailed calculations behind the physician profiling model as implemented at Health Clinic along with the reliability estimates using calculations suggested by Adams (2009) and Scholle et al. (2008). The chapter concludes with a discussion of future research topics.

3.2 Physician Profiling Model: Detailed Calculations

The physician profiling model described in section 2.4.2, and shown in Figure 2.2, is a hierarchical model with individual measures grouped into domains of care. The domains are scored across individual measures to create composites. The composites are then added together using weights determined by the organization. Finally, the weighted composite score is standardized resulting in a final ranking across physicians. The detailed model calculations start with the individual measures.

3.2.1 Individual Measures

All of the individual measures used in this model except for Generic Prescribing followed the HEDIS guidelines (NCQA, 2007). For each measure, a physician rate was determined by dividing the number of patients satisfying the numerator criteria (quality hits) by the total eligible population of patients for that physician (quality events). For the case study, a standardized Z

score was then calculated for each physician within each measure using the mean and standard deviation across all physician proportions as shown in the general Z score equation (Petruccelli et al., 1999) below.

$$Z_{ij} = \frac{X_{ij} - \mu_i}{\sigma_i}$$ Equation 3.1

Where,

X_{ij} = performance for measure i, physician j;

μ_i = average performance across physicians for measure i;

σ_i = standard deviation of performance across physicians for measure i

A Z score represents the number of standard deviations that a value is from the mean (Petruccelli et al., 1999), and allows comparison across values that inherently have different means and standard deviations. The standardized Z score calculation can be found in many applications. For instance, percentile ranks associated with standardized tests (e.g., GMAT, SAT, GRE) are calculated using Z scores (Johnson and Christensen, 2008). Z scores can also be found in the ranking of professionals. In the sport of golf, one study added Z scores from driving accuracy and driving distance to form a small composite Z score called "driving performance". This methodology adjusted the Professional Golfers' Association (PGA) ranking system moving Tiger Woods from 83rd in the PGA Tour Rankings to 1st using the new methodology (Wiseman et al., 2007). Another study ranked professional pitchers by calculating a weighted composite Z score using strikeout, walk, and groundball rates (Lederer, 2009). The weights were determined using regression of these rates on predicting earned run average (ERA) and runs allowed (RA) outcomes. We used the standardized Z scores for the individual measures to form composites.

3.2.2 Composite Measures

The individual measures were grouped into domains of quality of care and efficiency. The methodology continues by using weighted averages of the Z scores across each physician's patient population in each domain as suggested by Scholle et al. (2008). This approach accounts for the difficulty in achieving each measure, its variability and the distribution of eligible patients for each measure across physicians, thereby accounting for differences in patient panel characteristics. In computing a composite score per physician, measures were only used if they had more than four quality events. The resulting score per physician in the domain was then standardized to get the final ranking for each physician in each of the domains. Standardizing is necessary for comparison across domains because the weighted score is no longer a normalized Z score with a mean of zero and a standard deviation of one, and does not represent the number of standard deviations from the mean. The domains were then added using weights determined by the case study site in alignment with their organizational policy. The model equation as implemented at case study site using weighted average Z scores across domains of care is:

$$Z_j = \sum_{k=1}^{n} \Omega_k \sum_{i=1}^{l} \left[\frac{\left[\sum_{i=1}^{l} \frac{n_{ijk}}{N_{jk}} \frac{(p_{ijk} - \mu_{ik})}{\sigma_{ik}} \right] - \overline{X_{ik}}}{s_{ik}} \right] \qquad \text{Equation 3.2}$$

Where,

Z_j = composite score for physician j

Ω_k = organizational weights for domains k = 1,2,3,...,n set by organization

n_{ijk} = total quality events for measure i and physician j and domain k

p_{ijk} = performance proportion for measure i, physician j, domain k

N_{jk} = total quality events across all measures for physician j in domain k

μ_{ik} = average performance across physicians for measure i in domain k

σ_{ik} = standard deviation of performance across physicians for measure i in domain k

$\overline{X_{ik}}$ = average weighted Z score across physicians for measure i in domain k

s_{ik} = standard deviation of weighted Z score performance across physicians for measure i in domain k

Measures = i = 1,2,3,...,l

Physicians = j =1,2,3,...,m

Domain = k = 1,2,3,..,n

 For the same reason as the weighted composite, the final weighted score across domains was then standardized to create a final ranking for all physicians across the organization. This also provided the case study site a mechanism to use a minimum payout threshold based on the number of standard deviations from the mean.

This methodology did not include case-mix adjustments for patient population or medical practice characteristics in the ranking or weighting of the model but these could be considered in future applications.

The model was implemented at Health Clinic using Microsoft Excel. An example of the model is displayed in Figure 3.1. The rows represent individual physician scores. One benefit of using composites is an increase in sample size as shown in column EH relative to the measure denominators in columns DW and EC. All primary care physicians did not get a score for every measure. For example, the blank cells in the cervical cancer screening and breast cancer screening columns in Figure 3.1 represent either a pediatrician who would not have eligible patients for those measures or a family physician or internal medicine physician who did not have quality events for those measures.

=IF(EI5="","",STANDARDIZE(EI5,EI$268,EI$269))

DV	DW	DX	DY	EA	EB	EC	ED	EE	EG	EH	EI	EJ
CervicalCS_RATE	CervicalCS_Denom	CervicalCS_Numer	Z Score	Weighted Z Score	BreastCS_RATE	BreastCS_Denom	BreastCS_Numer	Z Score	Weighted Z Score	QOC Domain Denominator, across all measures with denominators >3	QOC Domain Weighted Z-Score	Final Physician Z-score for QOC Domain
					87.50%	16	14	0.627	0.101	99	0.585	1.148
89.80%	98	88	0.388	0.110	76.39%	72	55	0.088	0.018	347	0.306	0.435
87.45%	255	223	0.260	0.071	89.24%	251	224	0.712	0.192	933	0.406	0.690
										224	0.642	1.295
85.71%	7	6	0.165	0.018	81.25%	16	13	0.324	0.079	66	0.140	0.008
72.22%	18	13	-0.573	-0.055	75.00%	16	12	0.021	0.002	186	0.095	-0.106
78.48%	79	62	-0.231	-0.049	78.38%	74	58	0.185	0.036	375	-0.214	-0.898
										109	0.648	1.308
86.13%	238	205	0.188	0.054	79.81%	213	170	0.254	0.065	830	0.195	0.151
78.26%	46	36	-0.243	-0.063	79.41%	34	27	0.235	0.045	176	0.204	0.173
92.19%	64	59	0.519	0.170	69.23%	39	27	-0.259	-0.052	196	-0.061	-0.507
72.46%	69	50	-0.560	-0.091	62.07%	58	36	-0.607	-0.083	424	-0.199	-0.858
										40	0.360	0.573
90.00%	110	99	0.400	0.072	83.67%	98	82	0.442	0.071	611	0.508	0.950
84.21%	38	32	0.083	0.014	60.61%	33	20	-0.678	-0.099	227	-0.527	-1.697
										302	0.873	1.884
94.44%	36	34	0.643	0.073	94.12%	34	32	0.948	0.102	315	0.366	0.588
91.19%	159	145	0.465	0.144	76.19%	105	80	0.078	0.016	515	0.012	-0.319
88.24%	17	15	0.303	0.060	86.36%	22	19	0.572	0.146	86	0.138	0.004
76.92%	13	10	-0.316	-0.103	77.78%	9	7	0.155	0.035	40	0.062	-0.191
93.10%	29	27	0.569	0.101	81.25%	32	26	0.324	0.064	163	0.441	0.779
										225	0.191	0.138
79.30%	256	203	-0.186	-0.040	76.57%	239	183	0.097	0.019	1186	0.117	-0.051
81.30%	123	100	-0.076	-0.015	76.00%	100	76	0.069	0.011	647	-0.167	-0.776
90.91%	22	20	0.449	0.056	86.67%	30	26	0.587	0.099	177	0.482	0.883
77.22%	158	122	-0.300	-0.059	78.18%	165	129	0.175	0.036	805	-0.080	-0.554
89.63%	135	121	0.379	0.068	89.31%	131	117	0.715	0.125	752	0.557	1.075
83.21%	137	114	0.028	0.010	65.31%	98	64	-0.450	-0.113	391	-0.278	-1.060
100.00%	4	4	0.947	0.058	80.00%	5	4	0.263	0.020	65	0.247	0.282
82.22%	45	37	-0.026	-0.005	88.89%	36	32	0.695	0.097	257	0.297	0.412
										62	0.245	0.278
91.78%	73	67	0.497	0.111	87.10%	62	54	0.608	0.115	327	0.502	0.935
71.43%	14	10	-0.617	-0.184	81.82%	11	9	0.351	0.082	47	-0.025	-0.413
92.71%	192	178	0.548	0.190	81.60%	125	102	0.341	0.077	555	0.085	-0.132
87.18%	39	34	0.245	0.096	74.19%	31	23	-0.019	-0.006	100	-0.097	-0.597
87.05%	193	168	0.238	0.050	85.63%	160	137	0.536	0.094	913	0.313	0.452

Figure 3.1 Excel Model at Case Study Site

3.3 Reliability

Reliability is a measure of the ability to differentiate physicians and is an appropriate measure when quality events vary between physicians. Traditional methods of evaluating profiling tools focus on sample size selection for improved precision using methods of confidence intervals or power using the ANOVA F-statistic (Oehlert, 2000). These approaches focus on increasing sample size to increase the precision of the estimator. While this is important and present in the denominator of reliability, these evaluation methods do not address the variation between physicians due to varying quality events. For this reason, reliability has recently blossomed in the medical literature as a metric to consider in physician profiling (Scholle et al., 2008; Krein et al., 2002; Nyweide et al., 2009).

Reliability is a measure of signal to noise or the R^2 statistic in regression as the fit of the observed values to the true values. Reliability has been traditionally used as a measurement of reproducibility between known or true values and observed values. The calculation calls for the variability of the differences between the observed values and the true values divided by the total variability which includes measurement error. Physician performance reliability as presented by Adams (2009) using a two-level hierarchical model requires the same equation but instead of using the variability of the differences from the observed values and the true values (true values of physician performance are not available), the variability of the differences between physicians is used as the signal. This physician-to-physician variance can be considered as the variance that would be obtained if the true values were known.

The reliability estimate is calculated using the Spearman-Brown formula (Shrout and Fleiss, 1979) by dividing the physician-to-physician variance by the total of the physician-to-physician variance and the within physician variance, as shown below.

$$r = \frac{\sigma^2{}_{p-p}}{(\sigma^2{}_{p-p} + \sigma^2{}_{error})}$$

Equation 3.3

Where,

$\sigma^2{}_{p-p}$ = physician-to-physician variation

$\sigma^2{}_{error}$ = within physician variance

Reliability estimates vary between 0 and 1, and a reliability of 0.70 or higher is typically considered acceptable for psychometric purposes (Nunnelly and Bernstein, 1994). The estimate can be interpreted as the percentage of the total variation that can be explained by physician to physician variation. A reliability of one indicates that all of the variation exists between physicians. A reliability of zero indicates that all of the variation exists within physicians (or error variance). High reliability estimates suggest that it is easy to tell physicians apart. Low reliability estimates suggest that it is difficult to be confident that the physicians are different given the sample sizes, proportions and physician to physician variation. This study uses variance calculations similar to Adams (2009) and Scholle et al. (2008) for both the individual measures and the composites.

3.3.1 Individual Measure Calculations

The reliability estimate for the individual measures requires estimates of the physician-to-physician variance and the within physician variance. For the physician-to-physician variance, simply analyzing the variation across all physicians for each measure is not appropriate due to the varying sample sizes from physician to physician along with the presence of small sample sizes. The physician-to-physician variance is determined using a beta-binomial model for each measure using a SAS macro (Wakeling, 2004). Conceptually, this is estimating what physician-

to-physician variation would be if we had very large sample sizes for each physician. The alpha

and beta values from the SAS output are then used in the beta-binomial variance calculation

(Adam, 2009; Lu and Fang, 2003) to estimate the physician-to-physician variance, using the

following formula:

$$\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$ Equation 3.4

Where,

σ^2 = the physician-to-physician variance

α, β = beta-binomial positive shape parameters

3.3.1.1 Details of Beta-Binomial model

The beta-binomial model assumes that the variation of performance per physician and per

measure follows a binomial distribution and the probability of quality hits (responses) is assumed

to vary among physicians according to a beta distribution (Lu and Fang, 2003):

$$P(y_{ij}) = \binom{n_{ij}}{y_{ij}} \frac{B(\alpha_i + y_{ij}, \beta_i + n_{ij} - y_{ij})}{B(\alpha_i, \beta_i)}$$ Equation 3.5

Where,

y_{ij} = quality hits (or successes) for physician j in measure (group) i = sum of all binary

   responses, y_{ijk}

n_{ij} = total quality events (trials) for physician j in measure (group) i

α_i, β_i = beta-binomial positive shape parameters in measure (group) i

$B(\alpha_i, \beta_i) = \Gamma(\alpha_i)\Gamma(\beta_i)/\Gamma(\alpha_i + \beta_i)$ Where $\Gamma(\cdot)$ is the gamma function, $\alpha_i > 0$, and $\beta_i > 0$.

Under the reparameterization, the beta-binomial mean for measure (group) i is:

$$\mu_i = \frac{\alpha_i}{\alpha_i + \beta_i}$$

And, $\Phi_i = \dfrac{1}{(\alpha_i + \beta_i + 1)}$ = intra-physician correlation parameters for measure (group) i

Therefore, the expected value of y_{ij} is $E(y_{ij}) = n_{ij}\mu_{ij}$ and the intra-physician correlation =

$corr(y_{ij}, y_{ij'}) = \Phi_i$, where y_{ij} and $y_{ij'}$ are the responses from two different physicians or

$j \neq j'$.

The total Physician-to-Physician variance for measure i is then estimated by what the

physician-to-physician variation would be if we had very large sample sizes for each physician.

$$\sigma_i^2 = \dfrac{\alpha_i \beta_i}{(\alpha_i + \beta_i)^2 (\alpha_i + \beta_i + 1)}$$

The beta-binomial variance was calculated per individual measure along with an

expected value for quality hits per physician.

3.3.1.2 Details of Within-Physician Variance

Next, the average binomial variation per physician is used for the within-physician

variation for the individual measures (Adams, 2009), calculated as:

$$\sigma^2_{error} = \sigma^2_{ijk} = \dfrac{p_{ijk}(1 - p_{ijk})}{n_{ijk}} \qquad\qquad \text{Equation 3.6}$$

Where,

y_{ijk} = total quality successes for measure i and physician j and domain k

n_{ijk} = total quality events for measure i and physician j and domain k

$p_{ijk} = y_{ijk} / n_{ijk}$ = performance for measure i, physician j, domain k

Measures = i = 1,2,3,…,l

Physicians = j =1,2,3,…,m

Domain = k = 1,2,3,..,n

Each physician is given a within-physician variance and consequently an individual reliability. For comparison across measures, a within-physician variance can be calculated per measure using an average performance rate and average number of quality events. Then, the reliability of the individual measures can be determined at an average proportion and average number of quality events across the primary care physicians using the Spearman-Brown formula (Shrout and Fleiss, 1979).

3.3.2 Individual Measure Reliability Example

The reliability per physician and measure is determined along with a typical reliability estimate per individual measure. We illustrate the calculations using the well adolescent visit (WAV) measure.

3.3.2.1 WAV Physician-to-Physician Variance

The physician to physician variance was calculated using a beta-binomial SAS macro (Wakeling, 2004). The SAS output for the beta-binomial distribution is shown in Figure 3.2.
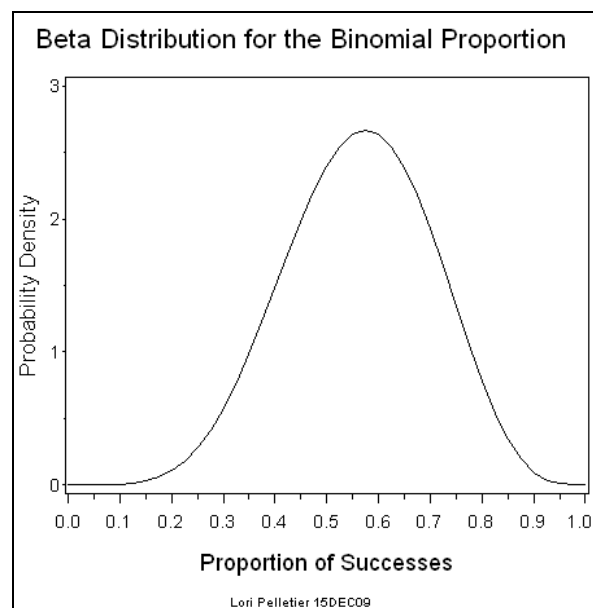
Figure 3.2 SAS Beta-Binomial Distribution for WAV measure

```
BETABIN Macro: Raw Data Summary

                                                        c
Item                           Value           Label2  Value2         nValue2

Data Set                       SASUSER.WAVFINAL                        .
Response Variable (Events)      Numer                   Numer          .
Response Variable (Trials)      Denom                   Denom          .
Number of Observations Read     207                                    .
Number of Observations Used     207                                    .
Number of Events                4436                                   .
Number of Trials                7034                                   .


BETABIN Macro: Simple Binomial Model


                    Standard              Prob
parameter   Estimate    Error    t Value   > |t|   Alpha     Lower      Upper

mu            0.6307    0.0058    109.59   <.0001    0.05    0.6194     0.6419
|mu-0.5|      0.1307    0.0058     22.70   <.0001    0.05    0.1194     0.1419


Lori Pelletier 15DEC09
BETABIN Macro: Beta-Binomial Model Parameters


                    Standard              Pr >
parameter   Estimate    Error    t Value    |t|    Alpha     Lower      Upper
N = 7034
mu            0.5623    0.01440    39.05   <.0001    0.05    0.5341     0.5905
alpha         6.4493    1.2728      5.07   <.0001    0.05    3.9542     8.9444
beta          5.0205    0.8983      5.59   <.0001    0.05    3.2596     6.7815
gamma         0.08019   0.01381     5.81   <.0001    0.05    0.05311    0.1073
theta         0.08719   0.01633     5.34   <.0001    0.05    0.05518    0.1192
|mu-0.5|      0.06228   0.01440     4.33   <.0001    0.05    0.03406    0.09051


BETABIN Macro: Variance-Covariance Matrix of Estimated Parameters

Label        mu       alpha       beta       gamma       theta
mu       0.000207   0.008221   0.002171   -0.00007    -0.00008
alpha    0.008221    1.6201     1.0935    -0.01745    -0.02063
beta     0.002171    1.0935     0.8069    -0.01222    -0.01445
gamma   -0.00007    -0.01745   -0.01222    0.000191    0.000226
theta   -0.00008    -0.02063   -0.01445    0.000226    0.000267
```

Figure 3.3 SAS Beta-Binomial Output for WAV measure

The alpha value (6.4493) and the beta value (5.0205) were input into the variance

calculation in equation 3.4, yielding a physician to physician variance estimate of 0.0197.

3.3.2.2 WAV Error Variance

The individual measure error variance is determined at the individual physician level and

at the average proportion and sample size across physicians using the formula suggested by

Adams (2009), and shown in equation 3.6. The average number of quality events for the Well

Adolescent Visit (WAV) measure is 34 and the average proportion is 0.524, corresponding to an

average error variance of 0.0073 as calculated in the Excel model and shown in Figure 3.4.

	Events, n	Hits	binomial p	Error Variance	a	b	mean, μ	theta,Φ	E(yij)	beta-binomial variance	Reliability
1	460	296	0.643478261	0.000498726	6.45	5.0205	0.562285	0.080194	258.6512	0.019737329	0.97535458
2	301	210	0.697674419	0.000700747	6.45	5.0205	0.562285	0.080194	169.2479	0.019737329	0.965713654
3	269	154	0.572490706	0.000909833	6.45	5.0205	0.562285	0.080194	151.2547	0.019737329	0.95593423
4	260	206	0.792307692	0.000632909	6.45	5.0205	0.562285	0.080194	146.1942	0.019737329	0.968929743
5	179	124	0.69273743	0.001189119	6.45	5.0205	0.562285	0.080194	100.6491	0.019737329	0.943176265
6	168	125	0.744047619	0.001133576	6.45	5.0205	0.562285	0.080194	94.46393	0.019737329	0.945686305
7	165	101	0.612121212	0.001438963	6.45	5.0205	0.562285	0.080194	92.77708	0.019737329	0.932048412
8	164	100	0.609756098	0.001450937	6.45	5.0205	0.562285	0.080194	92.21479	0.019737329	0.931521691
9	156	114	0.730769231	0.00126119	6.45	5.0205	0.562285	0.080194	87.71651	0.019737329	0.939939119
10	149	116	0.77852349	0.001157213	6.45	5.0205	0.562285	0.080194	83.78051	0.019737329	0.944616514
11	143	104	0.727272727	0.001387043	6.45	5.0205	0.562285	0.080194	80.4068	0.019737329	0.93433922
12	143	88	0.615384615	0.00165515	6.45	5.0205	0.562285	0.080194	80.4068	0.019737329	0.922629367
13	129	100	0.775193798	0.001350918	6.45	5.0205	0.562285	0.080194	72.5348	0.019737329	0.935939783
Avg n and avg p	33.98068	17.81239388	0.524191858	0.0073399	6.45	5.0205	0.562285	0.080194	116.3931	0.019737329	**0.728927216**
											Typical Experience

Figure 3.4 Reliability estimates for WAV in Excel

3.3.2.3 WAV Reliability Calculation

The reliability calculations were performed in Excel resulting in a reliability estimate for each

physician using equation 3.3. Sample results are shown in Figure 3.4. The resulting reliability

estimate at the average proportion and quality events per PCP was 0.729 using the Spearman-

Brown formula. The distribution of quality events per PCP and reliability estimates is shown in

Figure 3.5. The average number of quality events was 34 with an average proportion of 0.524

and physician to physician variance of 0.0197. Figure 3.5 indicates that reliability estimates

above 0.8 require sample sizes greater than 40. In order to address this issue of sample sizes, we

pool individual measures into composites.
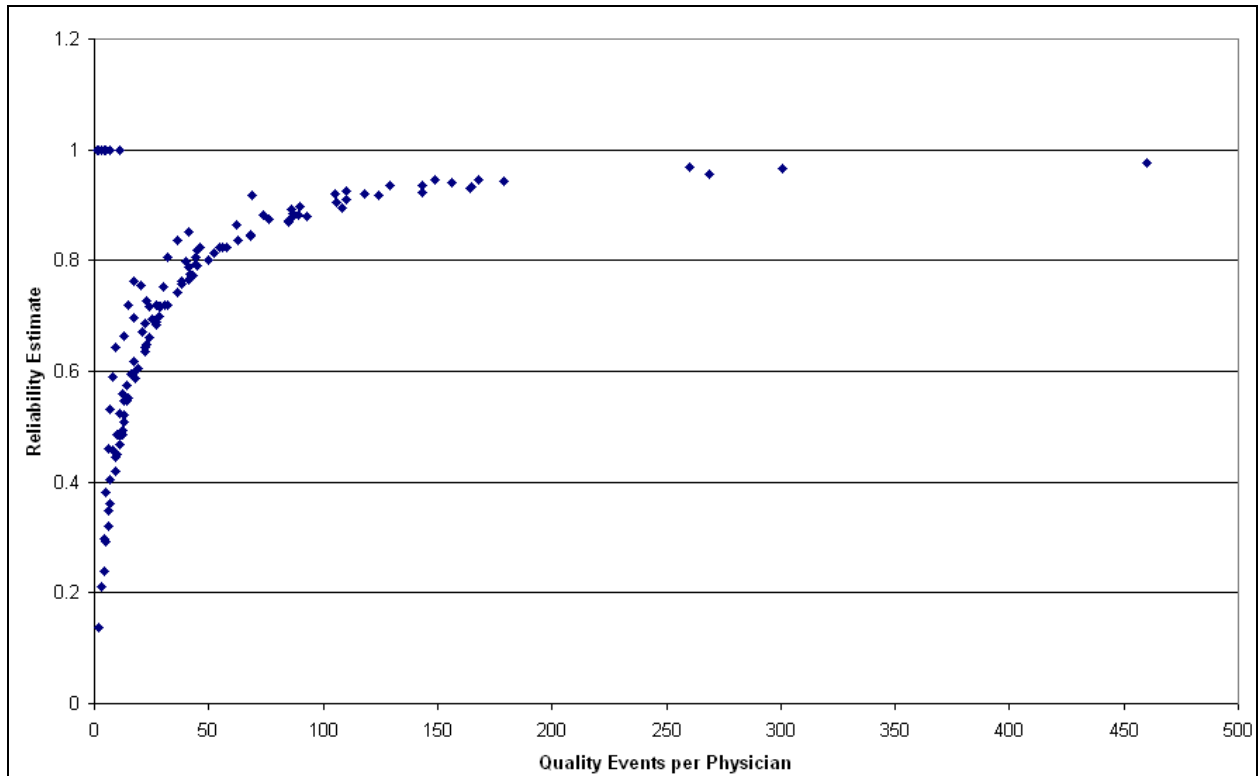
Figure 3.5 Reliability Estimates of WAV by the Number of Quality Events per PCP

3.3.3 Composite Measure Calculations

The reliability estimate for the composite measures requires estimates of the physician-to-physician variance and the within physician variance. We start with the within physician variance (error variances) which are then used to determine the physician to physician variance.

3.3.3.1 Details of Composite Error Variance

The composite error variances per physician are calculated as the summation of the squared weights multiplied by the individual measure average error variance. Adams (2009) suggested the following equation for physician j in domain k:

$$\sigma^2_{error, jk} = \sum_{i=1}^{l} (\frac{n_{ijk}}{N_{jk}})^2 (\frac{p_{ijk}(1-p_{ijk})}{n_{ijk}})$$

Equation 3.7

Where,

N_{jk} = total quality events across all measures for physician j in domain k

y_{ijk} = total quality successes for measure i and physician j and domain k

n_{ijk} = total quality events for measure i and physician j and domain k

$p_{ijk} = \frac{y_{ijk}}{n_{ijk}}$ = performance for measure i, physician j, domain k

Measures = $i = 1,2,3,\ldots,l$

Physicians = $j = 1,2,3,\ldots,m$

Domain = $k = 1,2,3,..,n$

An error variance was calculated per PCP and is based on the quality event distribution and the rates of the individual measures for that PCP. These error variance estimates are then used to determine the physician-to-physician variances.

3.3.3.2 Details of Composite Physician to Physician Variance

The physician-to-physician variance is determined using hierarchical linear modeling (HLM) for each composite using a PROC MIXED SAS procedure (Adams, 2009; Wakeling, 2004). A SAS input table was created for each measure and included the individual physician composite rates and their individual error variances. The type of HLM is called the one way random effects model with the following decomposition (Petruccelli, 1999):

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

Equation 3.8

Where,

y_{ij} = weighted Z score value for treatment effect i and physician j

μ = the average of the composite scores across all physicians

α_i = treatment effect, i of physician j (based on practice style, training, etc)

ε_{ij} = random error for treatment effect i and physician j

   The variance of y_{ij} is $\sigma_\alpha^2 + \sigma^2$

Where,

σ_α^2 = the physician-to-physician variance

σ^2 = the within physician variance or the measurement error.

   These terms are called components of variance or variance components. Thus the random effects model is sometimes called components of variance model. In order to obtain the physician to physician variance, we input the weighted z scores for each physician and the error variances for each physician into the PROC MIXED program.

   The physician-to-physician variance and error variance estimates are now used to calculate reliability estimates using the Spearman-Brown Formula. Because the error variances are determined per physician and represent the distribution of events for that particular physician, it is not appropriate to calculate a population reliability estimate at an average sample size and average proportion. It is more appropriate to simply calculate a median reliability across the individual PCP calculated reliabilities.

3.3.4 Composite Measure Reliability Example

   For each composite measure, a median reliability was calculated reflecting the combination of factors including sample size, physician-to-physician variance and measurement error necessary to achieve a reliability of 0.70. Measures were added to increase the total sample size per physician and variance such that physicians with a panel size greater than 30 could be differentiated using a composite score methodology. The assumption in equations 3.3 and 3.7 is that as the number of measures in the composite measure increases, the sample size increases, causing the error variance to decrease and the reliability to increase. This may not always be the

case. For instance, if a measure is added that has a high average proportion with low physician-to-physician variation or if a measure is added with a very high measurement error, the reliability of the composite could actually decrease. This phenomenon is demonstrated in section 3.4.1.

The individual physician reliabilities and median reliability per composite measure were determined using equations 3.3, 3.7 and 3.8. The calculations are illustrated using the quality of care composite measure.

3.3.4.1 Quality of Care Composite Error Variance

The composite error variance was calculated per primary care physician (PCP) as the sum of the squared weights multiplied by the individual measure average error variance. This calculation was performed in Excel as shown in Figure 3.6.

fx	=IF(G2="",0,(G2/FD2)^2*(L2))+IF(N2="",0,(N2/FD2)^2*(S2))+IF(U2="",0,(U2/FD2)^2*(Z2))+IF(AB2="",0,(AB2/FD2)^2*(AG2))+IF(AI2="",0,(AI2/FD2)^2*(AN2))+IF(AP2="",0,(AP2/FD2)^2*(AU2))+
X	IF(AW2="",0,(AW2/FD2)^2*(BB2))+IF(BD2="",0,(BD2/FD2)^2*(BI2))+IF(BK2="",0,(BK2/FD2)^2*(BP2))+IF(BR2="",0,(BR2/FD2)^2*(BW2))+IF(BY2="",0,(BY2/FD2)^2*(CD2))+IF(CF2="",0,(CF2/
	FD2)^2*(CK2))+IF(CM2="",0,(CM2/FD2)^2*(CR2))+IF(CT2="",0,(CT2/FD2)^2*(CY2))+IF(DA2="",0,(DA2/FD2)^2*(DF2))+IF(DH2="",0,(DH2/FD2)^2*(DM2))+IF(DO2="",0,(DO2/FD2)^2*(DT2))+IF(
	DV2="",0,(DV2/FD2)^2*(EA2))+IF(EC2="",0,(EC2/FD2)^2*(EH2))+IF(EJ2="",0,(EJ2/FD2)^2*(EO2))+IF(EQ2="",0,(EQ2/FD2)^2*(EV2))+IF(EX2="",0,(EX2/FD2)^2*(FC2))

CS_n	BreastCS_Numer	Z Score	Rank Across All PCPs	Weighted Z Score	Error Variance, p(1-p)/n	Total Denominators	Total Numerators	Composite Proportion	Error Variance, p(1-p)/n	across all measures with denominators >3	QOC Domain Weighted Z-Score	Physician Z-score for QOC Domain	Rank All QOC
14	13	0.887	90.17%	0.129	0.005	101.000	80.000	0.792	0.00118332	96	0.511	0.958	8
						94.000	63.000	0.670	0.00213524	94	0.104	-0.083	4
2	2	1.234	100.00%		0.000	83.000	66.000	0.795	0.00160791	80	0.346	0.537	6
16	14	0.627	83.82%	0.101	0.007	109.000	89.000	0.817	0.00113543	99	0.585	1.148	9
72	55	0.088	39.31%	0.018	0.003	351.000	281.000	0.801	0.00040821	347	0.306	0.435	6
251	224	0.712	86.13%	0.192	0.000	944.000	759.000	0.804	0.00015302	933	0.406	0.690	7
						232.000	175.000	0.754	0.00072861	224	0.642	1.295	9
16	13	0.324	54.34%	0.079	0.010	80.000	56.000	0.700	0.00202456	66	0.140	0.008	4
16	12	0.021	33.53%	0.002	0.012	189.000	141.000	0.746	0.00090029	186	0.095	-0.106	4
						40.000	27.000	0.675	0.00468750	38	0.180	0.110	5
						86.000	67.000	0.779	0.00191152	79	0.704	1.452	9
						22.000	16.000	0.727	0.00550964	20	0.077	-0.152	4
74	58	0.185	46.24%	0.036	0.002	377.000	253.000	0.671	0.00052411	375	-0.214	-0.898	1

Figure 3.6 Quality of Care Example of Error Variance in Excel

3.3.4.2 Quality of Care Composite Physician-to-Physician Variance

Physician-Physician variance for the composite was calculated using a PROC MIXED procedure in SAS as suggested by Adams (2009). An input table was created including the individual physician composite rate and their individual error variance. The SAS code shown in Figure 3.7 was used. The SAS input table is shown in Figure 3.8 and the SAS output is shown in

Figure 3.9. The residual variance bolded, 0.02814, denotes the physician-to-physician variance in

Figure 3.9.

```
data gdata;
set SASUSER.QOCVARIANCE;
col = _n_;
row = _n_;
value = errorvariance;
keep col row value;
run;
proc mixed data=SASUSER.QOCVARIANCE METHOD=REML ;
class ID;
model score =;
random ID / gdata=gdata ;
run;
```

Figure 3.7 SAS PROC MIXED Code for Quality of Care Composite

	A	B	C
1	ID	Score	Errorvariance
2	1	0.792079208	0.0011833184733
3	2	0.670212766	0.0021352416335
4	3	0.795180723	0.0016079145125
5	4	0.816513761	0.0011354263109
6	5	0.800569801	0.0004082084020
7	6	0.804025424	0.0001530192008
8	7	0.754310345	0.0007286067947
9	8	0.7	0.0020245640159
10	9	0.746031746	0.0009002945127
11	10	0.675	0.0046875000000
12	11	0.779069767	0.0019115243927
13	12	0.727272727	0.0055096418733
14	13	0.671087533	0.0005241099853
15	14	0.662420382	0.0010292724778
16	15	0.833333333	0.0052951388889
17	16	0.687150838	0.0010265844897
18	17	0.736842105	0.0069252077562
19	18	0.785714286	0.0011001947368
20	19	0.866071429	0.0007680814978
21	20	0.760998811	0.0002015267269

Figure 3.8 Excel Input Table for use in SAS for Quality of Care Composite

```
The SAS System        12:24 Tuesday, January 5, 2010   4

                        The Mixed Procedure

                        Model Information

              Data Set                   SASUSER.QOCVARIANCE
              Dependent Variable         Score
              Covariance Structure       Variance Components
              Estimation Method          REML
              Residual Variance Method   Parameter
              Fixed Effects SE Method    Model-Based
              Degrees of Freedom Method  Containment

The SAS System        12:24 Tuesday, January 5, 2010   5

                        The Mixed Procedure
                      Class Level Information

          Class    Levels   Values

          ID         264     1 2 3 4 5 6 7 8 9 10 11 12 13
                            14 15 16 17 18 19 20 21 22 23
                            24 25 26 27 28 29 30 31 32 33
                            34 35 36 37 38 39 40 41 42 43
                            44 45 46 47 48 49 50 51 52 53
                            54 55 56 57 58 59 60 61 62 63
                            64 65 66 67 68 69 70 71 72 73
                            74 75 76 77 78 79 80 81 82 83
                            84 85 86 87 88 89 90 91 92 93
                            94 95 96 97 98 99 100 101 102
                            103 104 105 106 107 108 109
                            110 111 112 113 114 115 116
                            117 118 119 120 121 122 123
                            124 125 126 127 128 129 130
                            131 132 133 134 135 136 137
                            138 139 140 141 142 143 144
                            145 146 147 148 149 150 151
                            152 153 154 155 156 157 158
                            159 160 161 162 163 164 165
                            166 167 168 169 170 171 172
                            173 174 175 176 177 178 179
                            180 181 182 183 184 185 186
                            187 188 189 190 191 192 193
                            194 195 196 197 198 199 200
                            201 202 203 204 205 206 207
                            208 209 210 211 212 213 214
                            215 216 217 218 219 220 221
                            222 223 224 225 226 227 228
                            229 230 231 232 233 234 235
                            236 237 238 239 240 241 242
                            243 244 245 246 247 248 249
                            250 251 252 253 254 255 256
                            257 258 259 260 261 262 263
                            264
```

Figure 3.9 SAS PROC MIXED Output for Quality of Care Composite

```
                            Dimensions
                  Covariance Parameters         1
                  Columns in X                  1
                  Columns in Z                264
                  Subjects                      1
                  Max Obs Per Subject         264

                        Lori Pelletier 05JAN10

The SAS System        12:24 Tuesday, January 5, 2010   6

                        The Mixed Procedure

                        Number of Observations

            Number of Observations Read          264
            Number of Observations Used          264
            Number of Observations Not Used        0

                        Iteration History

         Iteration    Evaluations    -2 Res Log Like      Criterion

              0            1        -190.51408601
              1            2        -180.90285825       0.00000004
              2            1        -180.90287121       0.00000000

                    Convergence criteria met.
Covariance Parameter
                          Estimates

                      Cov Parm      Estimate

                      Residual      0.02814

                        Fit Statistics

            -2 Res Log Likelihood          -180.9
            AIC (smaller is better)        -178.9
            AICC (smaller is better)       -178.9
            BIC (smaller is better)        -175.3

                        Lori Pelletier 05JAN10
```

Figure 3.9 Continued: SAS PROC MIXED Output for Quality of Care Composite

3.3.4.3 Quality of Care Reliability Calculation

The reliability calculation for the quality of care composite were performed in Excel

using equation 3.3 as shown in Figure 3.10. The median reliability experienced by 264

physicians was 0.978 using the Spearman-Brown Formula. The distribution of quality events per

PCP and reliability estimates are shown in Figure 3.11. The average number of quality events

was 252 with an average proportion of 0.72 and physician-to-physician variance of 0.0281. All

physicians who received a payout had a reliability greater than 0.70.

Events, n	Hits	Variance	binomial p	Proc Mixed Variance for p-p	Reliability
912.000	768.000	0.00013771	0.842105263	0.02814	0.995130081
1054.000	843.000	0.000142911	0.799810247	0.02814	0.994947078
819.000	685.000	0.000147982	0.836385836	0.02814	0.994768743
1058.000	807.000	0.000149664	0.762759924	0.02814	0.994709583
916.000	735.000	0.000152623	0.802401747	0.02814	0.994605556
1189.000	884.000	0.000152923	0.743481918	0.02814	0.994594995
944.000	759.000	0.000153019	0.804025424	0.02814	0.994591627
1030.000	778.000	0.00016304	0.755339806	0.02814	0.994239498
179.000	122.000	0.000987846	0.681564246	0.02814	0.966085866
167.000	119.000	0.000999014	0.71257485	0.02814	0.965715579
161.000	110.000	0.001008987	0.683229814	0.02814	0.965385184
166.000	129.000	0.001013974	0.777108434	0.02814	0.965220041
179.000	123.000	0.001026584	0.687150838	0.02814	0.964802718
					0.978420121
					Median of all PCP's reliability

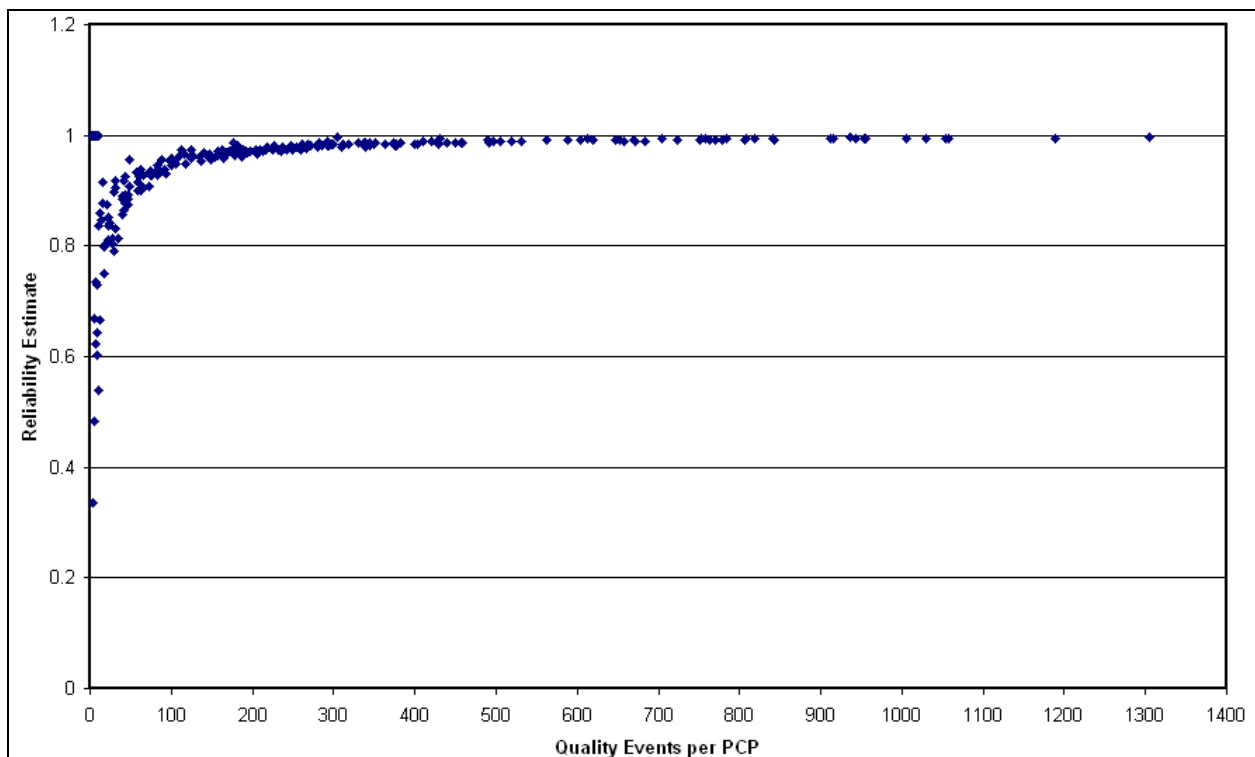Figure 3.10 Reliability estimates for Quality of Care in Excel

Figure 3.11 Reliability Estimates of Quality of Care Composite by the Number of Quality Events per PCP

3.4 Future Study Discussion

Two studies are framed for future consideration to help support the use of physician profiling in a health care organization. We discuss each in turn.

3.4.1 Reliability Sensitivities

Reliability measures the ability to differentiate one individual physician from another and is influenced by the physician-to-physician variance, sample size and performance rate of the measure. In our study, a composite reliability estimate was calculated for each individual primary care physician, which represented the ability to differentiate that physician from all the other physicians at that physician's performance rate and number of quality events. Ultimately, it would be desirable for every physician to have a reliability estimate greater than 0.70, but is it necessary? Or, is having the median reliability greater than 0.70 across physicians appropriate? A future study might explore this question by examining the sensitivity of sample size, physician-to-physician variance and performance rate of the measures on the reliability estimates and suggestions for the disposition of individual scores falling below the reliability threshold.

To frame this study, we conducted a sensitivity analysis using generic data to reveal some considerations regarding physician performance rate, sample size and physician-to-physician variation when reporting reliability estimates. In the reliability equation, the within-physician variance is the precision component which is dependent on sample size or quality events per physician and the performance rate for the measure, denoted by p. Establishing a minimum sample size to achieve a desired reliability estimate can be misleading as shown in Table 3.1. The results show that for a fixed sample size of 100 and a physician-to-physician variance of 0.002, reliability can vary with the performance rate. The error variance increases when the proportion is closest to 0.5 and hence reliability decreases. The results in Table 3.1 show that in order to maintain a target reliability at a fixed physician-to-physician variance, sample size would have to be adjusted according to the performance rate of the physician.

Table 3.1 Reliability versus physician rate

p	n	$\sigma^2 error$	$\sigma^2 p-p$	Reliability
0	100	0.0000	0.002	1.000
0.1	100	0.0009	0.002	0.690
0.2	100	0.0016	0.002	0.556
0.3	100	0.0021	0.002	0.488
0.4	100	0.0024	0.002	0.455
0.5	100	0.0025	0.002	0.444
0.6	100	0.0024	0.002	0.455
0.7	100	0.0021	0.002	0.488
0.8	100	0.0016	0.002	0.556
0.9	100	0.0009	0.002	0.690
1	100	0.0000	0.002	1.000

Reliability is also influenced by sample size and physician-to-physician variance. Table 3.2 shows an increase in reliability as sample size is increased with a constant performance rate and constant physician-to-physician variance. And, at the same performance rate and sample size, reliability increases as the physician-to-physician variance increases, typically due to the physician population becoming less homogeneous.

Table 3.2 Reliability versus sample size with constant rate and two levels of σ^2_{p-p}

p	n	$\sigma^2 error$	$\sigma^2 p-p$	Reliability
0.5	100	0.0025	0.002	0.444
0.5	50	0.0050	0.002	0.286
0.5	25	0.0100	0.002	0.167
0.5	10	0.0250	0.002	0.074
0.5	5	0.0500	0.002	0.038
0.5	100	0.0025	0.03	0.923
0.5	50	0.0050	0.03	0.857
0.5	25	0.0100	0.03	0.750
0.5	10	0.0250	0.03	0.545
0.5	5	0.0500	0.03	0.375

This sensitivity analysis is further compounded for composite calculations that use a weighted average across individual measure error variances. For simplicity, the above scenarios should be considered for individual measures or composites based on one measure.

Understanding the effects of sample size, physician performance rate and physician-to-physician variation on reliability, the question regarding the proper reliability estimate describing the population of physicians or the individual physicians can be revisited. Since we are

interested in the ability to differentiate the performance of one physician from another, individual reliability estimates should be of concern. Adams (2009) suggests 0.90 for an individual reliability target and 0.70-0.80 for a group reliability target. If a physician has a reliability estimate less than 0.90, the ability to differentiate performance is reduced and the probability of misclassification is increased. With the variation in reliability estimates caused by physician performance rates, sample sizes and physician-to-physician variances shown in Table 3.1 and 3.2, a methodology to capture the sensitivity of reliability estimates to minimize misclassification and a disposition plan for individual scores with reliabilities below the threshold should be explored.

3.4.2 Z Score Ranges

In the model used at Health Clinic, described in chapter 2, decision makers chose to standardize the individual measures and add across measures using a weighted average approach when creating composites. A Z score represents the number of standard deviations that a value is from the mean (Petruccelli et al., 1999), and allows comparison across values that inherently have different means and standard deviations The weighted Z score approach accounts for the difficulty in achieving each measure, its variability and the distribution of eligible patients for each measure across physicians, thereby accounting for differences in patient panel characteristics. We found that the Z score range (maximum Z score minus minimum Z score) across physicians per measure was highly influenced by the minimum quality event threshold and the number of physicians profiled. Our sensitivity analysis is shown in Table 3.3. For instance, for the Chlamydia (ages 16-20) measure shown in Table 3.3, simply removing the physicians who had one quality event in the denominator reduced the standard deviation from 35.86% to 25.82% and consequently, the Z score range increased from 2.79 to 3.87. By

definition of a Z score (Petruccelli, 1999) and shown in equation 3.1, as the variation in an

individual measure increases, the Z score range becomes compressed. Most of the variation due

to small sample sizes was reduced by simply excluding quality events equal to 1, as seen across

all measures in Table 3.3.

Table 3.3 Z Score Ranges across Quality Event thresholds for Select Measures

	ACE/ARBs				Chlamydia 16-20				Breast Cancer Screening				Diabetes A1C Testing (2x/year)			
	Rate	Denom	Numer	Z Score	Rate	Denom	Numer	Z Score	Rate	Denom	Numer	Z Score	Rate	Denom	Numer	Z Score
Average, All Quality Events	77.53%	21.13	16.54	0.00	46.90%	6.06	2.74	0.00	74.58%	61.46	49.44	0.00	68.90%	13.63	9.83	(0.00)
Standard Deviation	21.55%	18.85	16.27	1.00	35.86%	7.90	3.69	1.00	20.60%	63.60	53.74	1.00	24.82%	10.93	8.61	1.00
Variation	4.64%				12.86%				4.24%				6.16%			
#PCPs	179	179	179	179	142	142	142	142	173	173	173	173	181	181	181	181
Z Score Range				4.641				2.79				4.854				4.029
>= 2 Quality Events	77.77%	23.10	18.09	0.00	44.84%	8.56	3.84	0.00	75.81%	67.62	54.41	0.00	70.99%	15.02	10.86	0.00
Standard Deviation	18.01%	18.62	16.24	1.00	25.82%	8.63	4.07	1.00	14.53%	63.61	53.99	1.00	19.05%	10.63	8.46	1.00
#PCPs	163	163	163	163	95	95	95	95	157	157	157	157	163	163	163	163
Z Score Range				5.55				3.87				6.88				5.25
>= 3 Quality Events	77.27%	24.19	18.93	0.00	47.32%	10.77	4.89	0.00	76.83%	70.23	56.54	0.00	71.38%	15.88	11.48	0.00
Standard Deviation	17.65%	18.45	16.22	1.00	20.59%	8.97	4.21	1.00	12.59%	63.48	53.97	1.00	16.85%	10.42	8.36	1.00
#PCPs	155	155	155	155	71	71	71	71	151	151	151	151	153	153	153	153
Z Score Range				5.67				4.86				6.35				5.19
>= 4 Quality Events	76.80%	25.05	19.58	0.00	47.11%	12.20	5.52	0.00	77.34%	72.05	58.03	0.00	71.18%	16.59	11.99	0.00
Standard Deviation	17.56%	18.31	16.20	1.00	19.39%	9.06	4.28	1.00	12.15%	63.34	53.92	1.00	16.95%	10.24	8.29	1.00
Variation	3.08%															
#PCPs	149	149	149	149	60	60	60	60	147	147	147	147	145	145	145	145
Z Score Range				5.694				5.156				6.586				5.163
>= 5 Quality Events	76.55%	26.24	20.51	0.00	45.35%	14.25	6.35	0.00	77.77%	74.45	59.98	0.00	71.18%	16.59	11.99	0.00
Standard Deviation	17.59%	18.10	16.17	1.00	18.82%	9.03	4.38	1.00	11.53%	63.12	53.83	1.00	16.95%	10.24	8.29	1.00
#PCPs	141	141	141	141	48	48	48	48	142	142	142	142	145	145	145	145
Z Score Range				5.687				5.313				6.936				5.163
>= 10 Quality Events	76.44%	29.58	23.11	0.00	44.45%	19.34	8.59	0.00	79.08%	79.04	63.77	0.00	71.20%	20.57	14.90	0.00
Standard Deviation	17.07%	17.59	16.15	1.00	12.51%	8.23	4.19	1.00	9.12%	62.62	53.55	1.00	14.82%	9.73	8.23	1.00
#PCPs	120	120	120	120	29	29	29	29	133	133	133	133	102	102	102	102
Z Score Range				5.857				4.450				5.060				5.522

Thus, we found that as the minimum threshold of quality events per physician increased,

the standard deviation decreased, the number of physicians ranked decreased, and the Z score

range initially increased. However, as the remaining number of physicians fall below a threshold

(varied by measure), the Z score range begins to decrease and the distribution of Z scores

become nonsymmetrical around the mean of zero (non-normal distribution). Adding Z scores

across measures that have varying Z score ranges may be problematic especially when ranking

across physician specialties. For instance, if a primary care specialty is predominately ranked by

measures that have Z scores with values negatively skewed, the weighted average Z score could

yield biased results compared to other specialties. A future study could investigate the sensitivity in ranking and probability of misclassification of the ranked scores at varying minimum quality event thresholds of individual measures across physician specialties.

In conclusion, two studies were framed for future consideration to support physician profiling in a health care organization. The Reliability Sensitivities study investigates the variation of reliability estimates, misclassification of reliability, and methodologies for the disposition of scores with individual reliabilities less than a desired threshold. The Z Score Range study investigates the limitations of using weighted Z scores across heterogeneous populations by examining varying minimum quality event thresholds of individual measures.

3.4 Conclusion

This chapter presented the detailed calculations used in the development of the physician profiling model at Health Clinic along with the reliability estimates from both the individual measures and the composite measures. Examples were presented along with programming code and output to provide researchers and practitioners with comprehensive methodologies.

Two future studies were then discussed, focused on reliability sensitivities and Z Score range. These studies will help advance the usefulness and practical application of physician performance profiling.

3.5 References

Adams, J. L. "The Reliability of Provider Profiling: A Tutorial," *Rand,* 2009.

Johnson, B., and L. B. Christensen. *Educational Research : Quantitative, Qualitative, and Mixed Approaches*, Sage Publications, Los Angeles, 2008.

Krein, S. L., T. P. Hofer, E. A. Kerr, and R. A. Hayward. "Whom should we Profile? Examining Diabetes Care Practice Variation among Primary Care Providers, Provider Groups, and Health Care Facilities," *Health Services Research* (37:5), Oct, 2002, pp. 1159-1180.

Lederer, R. "Using Z-Scores to Rank Pitchers," *Baseball Analysts,* February 20, 2009, Available from: http://baseballanalysts.com/archives/2009/02/using_zscores_t.php.

Lü, Y., and J. Fang. *Advanced Medical Statistics*, World Scientific, Singapore; River Edge, NJ, 2003.

National Committee for Quality Assurance. "HEDIS 2007 Technical Specifications for Physician Measurement." 2007.

Nunnally, J. C., and I. H. Bernstein. *Psychometric Theory*, McGraw-Hill, New York, 1994.

Nyweide, D. J., W. B. Weeks, D. J. Gottlieb, L. P. Casalino, and E. S. Fisher. "Relationship of Primary Care Physicians' Patient Caseload with Measurement of Quality and Cost Performance," *JAMA : The Journal of the American Medical Association* (302:22), Dec 9, 2009, pp. 2444-2450.

Oehlert, G. W. *A First Course in Design and Analysis of Experiments*, W.H. Freeman, New York, 2000.

Petruccelli, J. D., B. Nandram, and M. Chen. *Applied Statistics for Engineers and Scientists*, Prentice Hall, Upper Saddle River, N.J., 1999.

Scholle, S. H., J. Roski, J. L. Adams, D. L. Dunn, E. A. Kerr, D. P. Dugan, and R. E. Jensen. "Benchmarking Physician Performance: Reliability of Individual and Composite Measures," *The American Journal of Managed Care* (14:12), Dec, 2008, pp. 833-838.

Shrout, P. E., and J. L. Fleiss. "Intraclass Correlations: Uses in Assessing Rater Reliability," *Psychological Bulletin* (86:2), 1979, pp. 420-428.

Wakeling, I. "BETABIN Macro.2004," Available at: http://www.sensory.org/library/files/SAS/betabin-v22.sas. Accessed May, 15, 2009.

Wiseman, F., Habibullah, M., Yilmaz, M. "A New Method for Ranking Total Driving Performance on the PGA Tour," *The Sport Journal* (10:1), 2007.

4.0 Examining the Relationships between Medical Home and Patient Experience

4.1 Introduction

The Patient-Centered Medical Home (PCMH) is an approach to providing comprehensive primary care for children, youth and adults that seeks to strengthen the physician-patient relationship by replacing episodic care based on illnesses and patient complaints with coordinated care and a long-term healing relationship. The PCMH model is a list of six principles recognized by physician associations and payers as a means of achieving the Institute of Medicine's vision of care that is safe, equitable, efficient, effectively, patient-centered and timely (IOM, 2001). The PCMH model is gaining national attention as a critical component in health care reform, given evidence that an emphasis on primary care can positively affect better outcomes (Schultz, 1995; Starfield et al., 2005; Kaiser Family Foundation, 2009). Plans for PCMH demonstration projects are under review as federal agencies such as the Centers for Medicare and Medicaid Services (CMS) and state agencies such as the Massachusetts Center for Health Policy and Research develop frameworks for evaluating PCMH feasibility, and the National Committee on Quality Assurance (Vesely, 2008) develops a recognition process for practices that adopt PCMH.

Physician Practice Connection – Patient-Centered Medical Home (PPC-PCMH) is a self-evaluation survey that can be used by practices to assess their level of medical homeness. The PPC-PCMH is endorsed by the National Committee for Quality Assurance (NCQA) for use in a three-level recognition program. Levels are determined by the presence of must-pass elements and the total points across elements. NCQA recognizes practices who achieved a level of recognition on the NCQA Web site and to other interested parties, including health plans.

The validity of the PPC-PCMH survey to characterize medical home characteristics is still under review, as is exploration of different levels of recognitions for predicting patient experience, quality of care and efficiency outcomes (Barr, 2008; Berenson et al., 2008).

Given the shortage of primary care physicians in many areas of the country (McGlynn et al., 2003), health care organizations are interested in attracting and retaining primary care physicians and in supporting their ability to coordinate care for patients. While PCMH holds promise of increasing quality of care in the US, there is little evidence to support such a claim (Rosenthal, 2008; Reid et al., 2009).

This study utilizes Pearson correlation coefficients to test four hypotheses about the relationships between constructs of the medical home and outcomes, using two sources of data from 16 practices that are part of a larger organization that we call Health Clinic. The characteristics of medical-homeness were measured through the PPC-PCMH Survey results, and related to patient experience data from Massachusetts Health Quality Partners (MHQP). As physician groups and individual physicians strive to deliver care that is patient-centered, it is increasingly important to receive feedback from patients to direct improvement areas. Studies such as this one that explore the relationships between medical home constructs, as defined by the PPC-PCMH survey, and measures of patient satisfaction is important because a truly patient-centered medical home is designed to enhance the patient experience as advocated by the American Academy of Family Physicians (http://www.aafp.org/online/en/home/membership/initiatives/pcmh.html, assessed April 12, 2010).

4.2 Literature Review

The literature review section is organized according to two components deemed critical for understanding the background and studies linking recognition programs to outcomes. We discuss each of these in turn.

4.2.1 Patient-Centered Medical Home and the PPC-PCMH Survey

The idea of a patient-centered medical home was conceptualized by the American Academy of Pediatrics (AAP) in 1967 as a central location for archiving a child's medical record. In 2002, the framework was expanded by the AAP to include operational characteristics such as accessible, continuous, comprehensive, family-centered, coordinated, compassionate and culturally effective care. By 2004, the American Academy of Family Physicians had developed their own medical home model. And then in 2006, the Advanced Medical Home: Joint Principles of the Patient-Centered Medical Home was issued by the American Academy of Family Physicians, the American College of Physicians, American Osteopathic Association and the American Academy of Pediatrics (Rosenthal, 2008). There are seven Joint Principles that form the basis of the PCMH model (www.ncqa.org , accessed May 28, 2009; Rogers, 2008):

1. *Personal physician.* Each patient has an ongoing relationship with a personal physician trained to provide first contact, continuous and comprehensive care.

2. *Physician directed medical practice.* The personal physician leads a team of individuals at the practice level who collectively take responsibility for the ongoing care of patients.

3. *Whole person orientation.* The personal physician is responsible for providing for all the patient's health care needs including care for all stages of life; acute care; chronic care; preventive services; and end of life care.

4. *Care is coordinated or integrated* across all elements of the complex health care system (e.g., subspecialty care, hospitals, home health agencies, nursing homes) and the patient's community (e.g., family, public and private community-based services).

5. *Quality and safety.* Practices advocate for their patients to support the attainment of optimal, patient-centered outcomes

6. *Enhanced access* to care is available through systems such as open scheduling, expanded hours and new options for communication between patients, their personal physician and practice staff.

7. *Payment* appropriately recognizes the added value provided to patients who have a patient-centered medical home.

Medical home is gaining popularity and in 2009, the Senate Finance Committee included PCMH as a model that might qualify for a wide-scale testing program of Chronic Care Management (Kaiser Family Foundation, 2009). Other state agencies are jumping on board with demonstration projects like the Massachusetts Center for Health Policy and Research Demonstration project. Rosenthal (2008) conducted a literature review to evaluate peer-reviewed literature on Medical Home for outcome summary and concluded that Medical Home has the ability to advance societal health.

Physician Practice Connection – Patient-Centered Medical Home (PPC-PCMH) is a self-evaluation survey (subject to audit) to assess the level of medical homeness in a practice. Formal recognition is then determined by submitting evidence to the National Committee for Quality Assurance (NCQA) which then reviews the evidence along with a possible site visit. NCQA may award Level 1, Level 2 or Level 3 Medical Home Recognition. If the practice attains at least Level 1, the NCQA disseminates information on the practice, its physicians and its level of performance to the NCQA Web site and to data users, including health plans and physician directory publishers (www.ncqa.org , accessed May 28, 2009).

The survey builds on the already established Physician Practice Connections (PPC) survey of the National Committee for Quality Assurance (NCQA), which together with the PPC-PCMH Recognition Program outlines specific structural and process standards that are

associated with excellent care. These standards include, for example, processes for access and communication, care coordination, patient tracking, test tracking and referral tracking. Nutting et al. (2009) urge the NCQA to incorporate new elements that focus on more patient-centeredness and engagement. Berenson et al. (2008) also suggest that current PPC-PCMH standards give too much weight to IT standards compared to access, communication, and care coordination. After studying thirty-six demonstration projects, Nutting et al. (2009) suggest recommendations for the transformation process to a medical home and indicate that one success factor is a practice's adaptive reserve, that is, its ability to change. The authors continue to caution primary care practices that the transformation is an on-going process and may go beyond the implementation of the PPC-PCMH elements and NCQA recognition to acquire additional organizational capabilities like leadership development, communication and adaptive reserve. The research question is whether the constructs of Medical homeness, as measured by the PPC-PCMH survey, are associated with improved patient experience and whether there are other confounding variables such as practice characteristics or measurement error limiting the ability to detect such correlations.

4.2.2 Linking Recognition Programs to Outcomes

Establishing whether recognition programs, either at the individual level or practice level, are correlated with better quality of care is important because the results can help support investment decisions for new models of delivering care. For example, Rosenthal and colleagues (2008) used claims data to assess whether physicians who were recognized by the Bridges to Excellence (BTE) program performed better than similar physicians on a standardized set of performance measures. BTE is a not-for-profit organization developed by employers, physicians, health care services, researchers, and other industry experts with the mission to create programs

that realign incentives around higher quality (www.bridgestoexcellence.org, accessed April 3, 2010) The BTE recognition program is self-reported (subject to audit) for certification and rewards physicians who met or exceeded performance criteria. Rosenthal and colleagues (2008) found that physicians who were BTE certified performed better than their nonrecognized colleagues on efficiency and quality measures including cervical cancer screening, breast cancer screening and A1C testing.

At the clinic recognition level, several studies have used comparative methodologies to examine structural characteristics and quality of care. Solberg et al. (2008) studied elements of the Chronic Care Model (Wagner, 1998) on diabetes care using Pearson correlations and found significant associations. Solberg and colleagues (2009) extended their study by conducting a cross-sectional survey of medical directors and found that organizational integration measures seem to be related to the presence of practice system components of the Chronic Care Model (1998). Landon et al. (2008) explored the relationship between characteristics of medical practices and quality measures for 1600 patients with coronary artery disease (CAD) from 225 medical practices and found that quality of care is not strongly influenced by characteristics of medical practices. Financial incentives, EHRs and care management approaches had little effect on measures.

Several studies have considered the relationship between patient experience and quality of care measures, using composite measures that combine several individual measures to increase sample size and improve reliability. Sequist et al. (2008) studied 373 practice sites and 119 individual primary care physicians (PCPs) in Massachusetts and found modest correlations between composites of prevention and Diabetes Mellitus and patient experience data from the Ambulatory Care Experiences Survey (ACES).   Additionally, Caldis (2007) found HEDIS

composite measures correlated to Consumer Assessment of Healthcare Providers and Systems (CAHPS®) survey components. While there are some studies that have examined the relationships between practice structures and quality of care (Landon et al., 2008; Solberg et al., 2009; Rosenthal et al., 2008) and relationships between quality of care and patient outcomes (Caldis, 2007; Sequist et al., 2008), there are very few studies that have explored the relationships between the practice structures and patient experience. This project extends the work in the literature by exploring the relationships between practice structures assessed by the PPC-PCMH tool and measures of patient experience.

4.3 Methodology

In this study, we used Pearson Correlation coefficients to study whether hypothesized correlations exist between two variables, patient experience and medical home elements. This study used data provided by Health Clinic, a private, non-profit integrated HC system in Massachusetts. The study population included sixteen practices in Health Clinic representing primary care delivery across pediatric primary care, family medicine and general internal medicine. Health Clinic provided two sources of data for this study, PPC-PCMH survey results and MHQP patient experience. Thirty-three practices were surveyed using the PPC-PCMH survey tool.   Twenty-nine practices received MHQP patient experience results. The sixteen practices included in both sources of data were used as the sample in this study. Other information included the number of physicians per practice and field notes from the PPC-PCMH surveys.

4.3.1 Data Sources

The first data source is the results of the PPC-PCMH survey instrument used to assess primary care practices for their level of medical homeness. The surveys started in late 2008 with

33 complete by mid-2009. Health Clinic assigned one surveyor, a nurse practitioner who met

with a team from each practice consisting of at least one physician, nurse, medical assistant,

practice manager and nonclinical staff. The surveys took approximately one hour to complete.

The assignment of one surveyor to administer the survey minimizes bias in how questions are

asked and responses are interpreted. However, measurement bias by way of incorrect

interpretation of the elements was possible. To minimize this possibility, three practices

received feedback from the NCQA recognition program and corrections in the raw survey data

for the remaining practices were made.

The surveyor entered the raw survey data into an excel workbook that computed Level

Recognition and segregated and analyzed the practices by size. The PPC-PCMH survey contains

9 standards, which consist of 30 elements overall, as shown in Table 4.1. For instance, within

element 1A (Access and Communication Processes in Table 4.1), there are 12 items to which the

practice answers yes or no. If the practice answers yes to 1-3 items, one point is awarded. If the

practice answers yes to 4-6 items, two points are awarded. If the practice answers yes to 7-9,

three points are awarded. And finally, if the practice answers yes to 10-12 items, the entire four

points for that element are awarded.

There are 10 must-pass elements, which along with the total score determine the level of

recognition given by NCQA. Must-pass elements are denoted by asterisks in Table 4.1, and

passing consists of scoring 50 percent or more, or greater than half of the points for that element.

To receive Level 1 recognition, practices must pass five of the required elements, as well as

receive a total score between 25 and 49. In order to receive Level 2 recognition, the practice

must pass all 10 must-pass elements and earn a total score between 50 and 74. Level 3

recognition requires the practice to pass all 10 must-pass elements and earn a total score of 75 or

more.

Table 4.1 PPC-PCMH Survey Standards and Elements

Standard	Element
1 Access and Communication (9 points)	1A: Access and Communication Processes*
	1B: Access and Communication Results*
2 Patient Tracking and Registry Functions (21 points)	2A: Basic System for Managing Patient Data
	2B: Electronic System for Clinical Data
	2C: Use of Electronic Clinical Data
	2D: Patient Tracking and Registry Functions: Organizing Clinical Data*
	2E: Patient Tracking and Registry Functions: Identifying Important Conditions*
	2F: Use of System for Population Management
3 Care Management (20 points)	3A: Care Management: Guidelines for Important Conditions*
	3B: Preventive Service Clinician Reminders
	3C: Practice Organization
	3D: Care Management for Important Conditions
	3E: Continuity of Care
4 Patient Self-Management Support (6 points)	4A: Documenting Communication Needs
	4B: Patient Self-Management Support*
5 Electronic Prescription (8 points)	5A: Electronic Prescription Writing
	5B: Prescribing Decision Support - Safety
	5C: Prescribing Decision Support - Efficiency
6 Test Tracking (13 points)	6A: Test Tracking and Follow-up*
	6B: Electronic System for Managing Tests
7 Referral Tracking (4 points)	7A: Referral Tracking*
8 Performance Reporting and Improvement (15 points)	8A: Performance Reporting and Improvement: Measures of Performance*
	8B: Patient Experience Data
	8C: Performance Reporting and Improvement: Reporting to Physicians*
	8D: Setting Goals and Taking Action
	8E: Reporting Standardized Measures
	8F: Electronic Reporting - External Entities
9 Advanced Electronic Communications (4 points)	9A: Availability of Interactive Web site
	9B: Electronic Patient Identification
	9C: Electronic Care Management Support

* Must Pass Elements

The second data source provided by Health Clinic is patient experience data extracted from the Massachusetts Health Quality Partners (MHQP) Patient Experience Survey conducted in late 2009. The MHQP Patient Experience Survey tool is a fifty question tool developed from two nationally recognized, validated surveys- the Ambulatory Care Experiences Survey (ACES) and the Consumer Assessment of Healthcare Providers and Systems (CAHPS) Survey. The ACES survey was developed by MHQP and researchers from Tufts New England Medical Center (http://160.109.101.132/icrhps/resprog/thi/aces.asp , Accessed April 19, 2010) and the CAHPS Survey was developed by the Agency for Healthcare Research and Quality (AHRQ) and endorsed by the national Quality Forum (NQF) (https://www.cahps.ahrq.gov/default.asp, accessed April 19, 2010). MHQP surveys a sample of patients from participating health care organizations at the practice level. Sample size is determined from reliability targets and non-response rates from previous years. Of a total of 21,000 surveys conducted by MHQP, 1500 surveys were complete for Health Clinic which equates to approximately100 surveys per practice. MHQP has eight measures within two domains, as shown in Table 4.2.

Table 4.2 MHQP Patient Experience Measures

Quality of Doctor-Patient Interaction

- Communication (How Well Doctors communicate with patients)

- Integration of Care (How Well Doctors coordinate care)

- Knowledge of patient (How Well Doctors know their Patients)

- Health Promotion (How Well Doctors give Preventive care and advice)

Organization/Structural Features of Care

- Organizational Access (Getting Timely Appointments, Care and Information)

- Visit based Continuity (Getting continuity of care across medical services)

- Clinical Team (Getting Quality Care from Other Doctors and Nurses in the office)

- Office Staff (Getting Quality Care from Staff in the doctor's office).

Patients are asked multiple questions for each measure and are required to answer each question with: never, almost never, sometimes, usually, almost always, or always. The results are case-mix adjusted according to age, gender, education, chronic conditions, race, language, health plan and region. The results represent the adjusted mean score across all of the questions in each measure. MHQP provides health care organizations with state averages per measure, reliability per measure, site and practice-type statistics, response frequencies per item and measure, confidence intervals, statistical comparisons between Health Clinic practices and demographics of respondents. The validity of the survey to measure patient experience has been extensively studied in the literature with good results (Safran et al., 2006; Rodriguez et al., 2006).

The patient questions and survey results are publicly communicated on the MHQP website (http://www.mhqp.org/default.asp?nav=010000, accessed Sept 21, 2009), if a minimum sample size of patient respondents for reliability is obtained. The publically reported results are

presented as the practice's score in percentile terms relative to all practices statewide and are grouped into four levels represented by stars.

MHQP requires that each practice have at least three active physicians with at least fifty patients in their panel across all five health plans in Massachusetts. Additionally, if the number of respondents was less than the minimum size determined for reliability, then the resulting scores were not reported publically but are reported privately as long as the sample size is greater than 16. As a result, Health Clinic received survey results on approximately 30 practices. The PPC-PCMH survey was completed on 33 practices. This study collected results from practices where both sources of data were available, yielding a sample size of 16 practices. All of the practices used in this study had acceptable reliability results for the MHQP survey.

4.3.2 Analysis

A matrix (shown in Table 4.3) was created by medical management at Health Clinic to facilitate hypothesis generation regarding predicted correlations between the structures identified by the PPC-PCMH survey and measures in the MHQP patient experience survey. Three plus signs in Table 4.3 represent a strong predicted association between the elements. Scatter plots were generated to explore the relationships between the bivariate variables. The mean, standard deviation and range were calculated for each PPC-PCMH element and each MHQP patient experience measure. Pearson correlation coefficient analysis was then used to determine if there was a significant correlation between elements of the PPC-PCMH and the measures used in the MHQP Patient Experience Survey. Because all of the measures are assumed to be interval-level data, and scatter plots suggest linear associations, we chose Pearson correlations over other methods of association (Petruccelli et al., 1999). The Pearson correlation coefficient (Petruccelli et al., 1999) is calculated as:

$$r = \frac{1}{n-1} \sum_{i=1}^{n} \left(\frac{X_i - \overline{X}}{s_x} \right) \left(\frac{Y_i - \overline{Y}}{s_Y} \right)$$

Equation 4.1

Where,

n = the number of practices

X_i = the value of the PCMH element for practice i

$\overline{X}$ = the average of the PCMH element across all practices

S_x = the standard deviation of the individual values of the PCMH element

Y_i = the value of the MHQP measure for practice i

$\overline{Y}$ = the average of the MHQP measure across all practices

S_Y = the standard deviation of the individual values of the MHQP measure

Table 4.3 Predicted Correlations for Hypothesis Generation

Correlation Matrix		MHQP: Quality of Doctor-Patient Interaction				MHQP: Organization/Structural Features of Care			
		How Well Doctors communicate with patients	How Well Doctors coordinate care	How Well Doctors know their Patients	How Well Doctors give Preventive care and advice	Getting Timely Appointments, Care and Information	Getting continuity of care across medical services	Getting Quality Care from Other Doctors and Nurses in the office	Getting Quality Care from Staff in the doctor's office
		Communication	Integration of Care	Knowledge of Patient	Health Promotion	Organizational Access	Visit-based continuity	Clinical Team	Office Staff
Medical Home Element	Element description								
1A	Access and Communication **Processes**					+++			
1B	Access and Communication **Results**					+++			
2A	Basic System for Managing Patient Data	+	+	+	+		+	+	+
2B	Electronic **System** for Clinical Data	+	+	+	+		+	+	
2C	**Use** of Electronic Clinical Data	+	+	+	+		+	+	
2D	Patient Tracking and Registry Functions: Organizing Clinical Data	+	+	+	+		+	+	
2E	Patient Tracking and Registry Functions: Identifying Important Conditions	+	+	+	+		+	+	
2F	Use of System for Population Management	+	+	+	+	+	+	+	+
3A	Care Management: Guidelines for Important Conditions	+	+	+	+		+	+	+
3B	Preventive Service Clinician Reminders	+	+	+	+++	+	+	+	+
3C	Practice Organization					+	+	+++	+
3D	Care Management for Important Conditions	+	+	+	+	+	+	+	+
3E	Continuity of Care	+	+	+	+	+	+	+	+
4A	Documenting Communication Needs	+	+	+	+	+	+	+	+
4B	Patient Self-Management Support	+	+	+	+	+	+	+	+
5A	Electronic Prescription Writing		+			+			
5B	Prescribing Decision Support - Safety		+			+			
5C	Prescribing Decision Support - Efficiency		+			+			
6A	Test Tracking and Follow-up	+	+	+	+	+	+	+	+
6B	Electronic System for Managing Tests	+	+	+	+	+	+	+	+
7A	Referral Tracking		+		+	+	+		
8A	Performance Reporting and Improvement: Measures of Performance		+		+	+			
8B	Patient Experience Data	+	+	+	+	+	+	+	+
8C	Performance Reporting and Improvement: Reporting to Physicians	+	+	+	+	+	+	+	+
8D	Setting Goals and Taking Action	+	+	+	+	+	+	+	+
8E	Reporting Standardized Measures		+		+	+			
8F	Electronic Reporting - External Entities		+		+	+			
9A	Availability of Interactive Web site	+	+	+	+	+	+	+	+
9B	Electronic Patient Identification	+	+	+	+	+	+	+	+
9C	Electronic Care Management Support	+	+	+	+	+	+	+	+

150

Scatter plots are necessary because a correlation can never by itself adequately summarize a set of bivariate data (Petruccelli et al., 1999). The correlation should be evaluated to determine whether it represents a real linear association (or another functional relationship, for example, quadratic) between the two variables or whether it is a reflection of sampling error. The data should be tested to determine whether convincing evidence is provided that the population correlation is different than zero. Hypothesis testing was conducted to determine the significance of the Pearson correlation coefficient using the statistical package SPSS. The following test statistic is used to test the null hypothesis that the Pearson coefficient equals zero (Petruccelli et al., 1999).

Test Statistic:

$$t^* = (r - \rho_0)\sqrt{\frac{n-2}{(1-r^2)(1-\rho_0^2)}}$$ Equation 4.2

P-values:

$$H_0 : \rho = \rho_0$$
$$H_a : \rho \neq \rho_0$$

Where, $$p\pm = 2 * P(t > |t^*|)$$

- the population correlation coefficient is ρ and we make inferences on it through the sample correlation coefficient r

- n is the sample of ordered pairs (X,Y)

- the sample is from a Bivariate Normal distribution

- r is the sample correlation coefficient

- the test statistic follows a Student's t distribution with degrees of freedom, v=n-2.

- the significance level of our tests, α, is equal to 0.05. The tests are said to give a statistically significant result if the p-value is smaller than 0.05.

The MHQP patient experience data is continuous data in the interval between 0 and 100. The PPC-PCMH data is discrete data except for the total score which is continuous data in the interval between 0 and 100. If both variables are normally distributed and independent, this implies they are jointly normally distributed or have a bivariate normal distribution. This assumption is weak due to a small sample size.

4.3.3 Hypotheses

Four hypotheses were formulated from the predicted correlations. We discuss these in turn.

4.3.3.1 MHQP Average and Total PPC-PCMH Score

The delivery of primary care that is patient-centered encompasses many facets of patient experience and practice structures. The overall experience a patient has with a practice involves a complex set of events, across multiple visits, from initiating an appointment to executing the patient plan at ancillary offices. Some systems that support primary care include appointments, test tracking, reminder systems, electronic health records, and electronic prescribing.

The **first hypothesis** is that if more elements of a medical home are in place, as measured by the PPC-PCMH, patients should be more satisfied. The test statistic examines whether the average MHQP Patient Satisfaction Score is correlated with the total score from the PPC-PCMH Score. The average MHQP score is used as a snapshot of the complexity of the patient experience. The total PPC-PCMH score is the sum of unique processes and results across practice structures.

4.3.3.2 Organizational Access

Primary care practices provide quality health care by giving patients appointments, care, and information in a timely way. The MHQP survey asks patients six questions on how well they receive timely appointments, care and information. The PPC-PCMH survey asks practices whether they have processes and results that ensure they are providing timely appointments, care and information.

The **second hypothesis** is if there are processes within a practice for patient access and communication, as well as results of those practices being in place, the patients should indicate that they are able to get an appointment or information when they need it. The test statistic examines whether the MHQP Patient Experience Results on Patient Access and Communication are associated with the Medical Home results on Patient Access and Communication (Elements 1A and 1B Must-Pass Elements). For this analysis, the individual results for elements 1A and 1B were added together to represent the intention of first PPC-PCMH standard. Table 4.4 shows the detailed questions for the PPC-PCMH element and MHQP patient experience measure. The first column lists MHQP patient questions for the Organizational Access Measure. The second column lists the PPC- PCMH survey questions for elements 1A (processes) and 1B (results).

Table 4.4 MHQP and MH Access and Communication Comparison

MHQP, Organizational Access: Getting Timely Appointments, Care and Information	PCMH: Access and Communication Processes (1A) and Results (1B)
For patients, MHQP asked: In the last 12 months... 1. When you called your (or your child's) personal doctor's office for an appointment you (or your child) needed right away, how often did you get an appointment as soon as you needed it? 2. When you scheduled an appointment for a check-up or routine care at your (or your child's) personal doctor's office, how often did you get it as soon as you needed it? 3. When you called your (or your child's) personal doctor's office with a medical question during regular office hours, how often did you get an answer to your question that same day? 4. When you called your (or your child's) personal doctor's office after regular office hours, how often did you get the help or advice you needed? 5. When you had an appointment at your (or your child's) personal doctor's office, how often were you (or your child) taken to the exam room within 15 minutes of your appointment time? 6. Once you (and your child) were in the exam room, how often did the person you were scheduled to see come in within 15 minutes?	For each practice, PCMH asked: The practice establishes in writing standards and has results for the following processes to support patient access: 1. Scheduling each patient with a personal clinician for continuity of care 2. Coordinating visits with multiple clinicians and/or diagnostic tests during one trip 3. Determining through triage how soon a patient needs to be seen 4. Maintaining the capacity to schedule patients the same day they call 5. Scheduling same day appointments based on practice's triage of patients' conditions 6. Scheduling same day appointments based on patient's/family's requests 7. Providing telephone advice on clinical issues during office hours by physician, nurse or other clinician within a specified time 8. Providing urgent phone response within a specific time, with clinician support available 24 hours a day, 7 days a week 9. Providing secure e-mail consultations with physician or other clinician on clinical issues, answering within a specified time 10. Providing an interactive practice Web site 11. Making language services available for patients with limited English proficiency 12. Identifying health insurance resources for patients/families without insurance

4.3.3.3 Health Promotion

Preventive care means taking actions to prevent health problems before they start, or finding problems early when they can treated. Preventive care includes cancer screening procedures like colonoscopies and mammograms, and immunizations to prevent infectious diseases. Preventive advice includes suggestions from a physician about ways patients can improve health or prevent problems, including lifestyle changes. For instance, a physician may recommend that a patient stop smoking, maintain a healthy weight, or exercise each day. The MHQP survey asks patients five questions regarding preventive care and preventive advice. The PPC-PCMH survey determines whether the practice has processes to identify patients by age, gender and status of preventive services, and to prompt the clinician about preventive services at the point of care.

The **third hypothesis** is that if there is paper-based or electronic system with guideline-based reminders for preventive services, patients should indicate that they received preventive care and advice. The test statistic examines whether the MHQP Patient Experience Results on Preventive Care and Advice are associated with the PPC-PCMH results on Preventive Service Clinician Reminders (Elements 3B). Table 4.5 shows the detailed questions for the PPC-PCMH element and MHQP patient experience measure. The first column lists the MHQP patient questions for the Preventive Care and Advice; the second column lists PCMH survey questions in element 3A.

Table 4.5 MHQP and MH Preventive Care Comparison

MHQP, Health Promotion: How Well Doctors Give Preventive Care and Advice	PCMH: 3B: Preventive Service Clinician Reminders
For adult patients, MHQP asked: In the last 12 months... 1. Did your personal doctor's office remind you to get preventive care that you were due to receive (for example, flu shot, cancer screening, mammogram, eye exam)? 2. Did you and your personal doctor talk about a healthy diet and healthy eating habits? 3. Did you and your personal doctor talk about the exercise or physical activity you get? 4. Did you and your personal doctor talk about things in your life that worry you or cause you stress? 5. Did your personal doctor ask whether there was a period of 2 weeks or more when you felt sad, empty, or depressed?	For each practice, PCMH asked: The practice uses a paper-based or electronic system with guideline-based reminders for the following services when seeing the patient: 1. Age-appropriate screening tests 2. Age-appropriate immunizations (e.g., influenza, pediatric) 3. Age-appropriate risk assessments (e.g., smoking, diet, depression) 4. Counseling (e.g., smoking cessation)?

4.3.3.4 Clinical Team

While physicians are responsible for directing and coordinating patient care, managing patient care is usually a team effort that involves all members of the practice who interact with patients (i.e., physicians, nurses, medical assistants, and nonclinical staff). The MHQP survey measures how well other physicians and nurses in a particular physician's office provided quality care. The PPC-PCMH survey measures the extent of clinical team involvement in providing care delivery services.

The **fourth hypothesis** considers that if there is a team approach in managing patient care, patients should indicate that they received quality care from other doctors

and nurses in the office. The test statistic examines whether the MHQP Clinical Team results are correlated with the PPC-PCMH results on Practice Organization (Elements 3C). Table 4.6 shows the detailed questions for the PPC-PCMH element and MHQP patient experience measure, with MHQP patient questions for the Clinical Team measure listed in the first column and PPC-PCMH survey questions for element 3C in the second column.

Table 4.6 MHQP and MH Clinical Team Comparison

MHQP, Clinical Team: Getting Quality Care from Other Doctors and Nurses in the Office	PCMH, 3C: Practice Organization
For patients, MHQP asked: In the last 12 months when you saw other doctors or nurses in your (or your child's) personal doctor's office... 1. How often did these other doctors and nurses explain things in ways that were easy to understand? 2. How often did you feel that these other doctors and nurse had all the information they needed to correctly diagnose and treat your (or your child's) health problems? 3. How often did these other doctors and nurses spend enough time with you (and your child)? 4. Overall, how would you rate the care you (or your child) got from these other doctors and nurses?	For each practice, PCMH asked: The care team manages patient care in the following ways: 1. Nonphysician staff remind patients of appointments and collect information prior to appointments 2. Nonphysician staff execute standing orders for medication refills, order tests and deliver routine preventive services 3. Nonphysician staff educate patients/families about managing conditions 4. Nonphysician staff coordinate care with external disease management or case management organizations.

4.4 Results

The main descriptive characteristics of the participating practices are reported in Table 4.7. The number of physicians in each practice ranged from 3 to 26 with a median of 7. Most of the practices (59%) were general internal medicine and 38% of the practices included full time residents.

Table 4.7 Description of Participating Practices (n=16)

Characteristic	No.	Percentage
No. of physicians		
3-8	8	50
>8	8	50
Type of practice		
Internal Medicine	10*	59
Family Medicine	5	29
Pediatric Primary Care	2	12
Residency		
Full Time Residents	6	38
Part Time Residents	1	6
No Residents	9	56

* One practice has both Internal Medicine and Pediatric Primary Care

Table 4.8 provides descriptive information on both the PPC-PCMH scores and the MHQP patient experience scores. Only the measures used in this study are reported. The average score across all practices was 72.8 (out of a 100) for the PPC-PCMH total score, with elements studied having average scores of 3.5 (out of 9) for access and communication, 3.6 (out of 4) for preventive service clinician reminders, and 2.6 (out of 3) for practice organization.

Table 4.8 also reports on the average MHQP patient experience score of 84.1 (out of 100) with a range between 75.7 and 88.6. The individual MHQP measures varied substantially as seen by the Clinical Team measure having the highest mean, with a value of 86.0 (out of 100) and a standard deviation of 5.5, and the Health Promotion measure with the lowest mean, equal to 62.4 (out of 100) and a standard deviation of 7.5.

Table 4.8 Practice Scores for PPC-PCMH and MHQP Patient Experience (n=16)

Measure	Mean (SD)	Range
PPC-PCMH Total Score	72.8 (14.5)	30.3-94.3
1A+1B: Access and Communication	3.5 (2.7)	0-7.8
3B: Preventive Service Clinician Reminders	3.6 (0.9)	1.0-4.0
3C: Practice Organization	2.6 (0.5)	1.5-3.0
MHQP Patient Experience Average Score	84.1 (3.1)	75.7-88.6
Health Promotion	62.4 (7.5)	54.4-86.6
Organizational Access	82.8 (4.2)	75.6-89.5
Clinical Team	86.0 (5.5)	77.9-96.2

Scatter plots of PPC-PCMH scores and the MHQP patient experience scores were created for the four hypotheses and Pearson correlation coefficients and hypotheses results are shown in Table 4.9.

Table 4.9 Pearson Correlations Between MHQP Patient Experience Scores and PPC-PCMH Scores Per Practice (n=16)

Hypothesis	MHQP	PPC-PCMH	Pearson Correlation	P value
1	MHQP Average Score	PPC-PCMH Total Score	-0.27	0.307
2	Organizational Access	Access and Communication	-0.72	0.002*
3	Health Promotion	Preventive Service Clinician Reminders	0.17	0.524
4	Clinical Team	Practice Organization	-0.49	0.053

*Significant at the 0.05 level

4.4.1 Hypothesis 1: MHQP Average and Total PPC-PCMH Score

The scatter plot of MHQP patient experience average scores (100 maximum) and the PPC-PCMH total scores (100 maximum) for each practice (Figure 4.1) and the calculated Pearson correlation coefficient of -0.27 (Table 4.9) indicate a very weak

negative linear association. The hypothesis results indicate that the MHQP patient experience average scores for each practice and the PPC-PCMH total scores for each practice are negatively correlated with a p-value of 0.307, not significant at the 0.05 level. That is, there is no significant correlation between the overall scores.
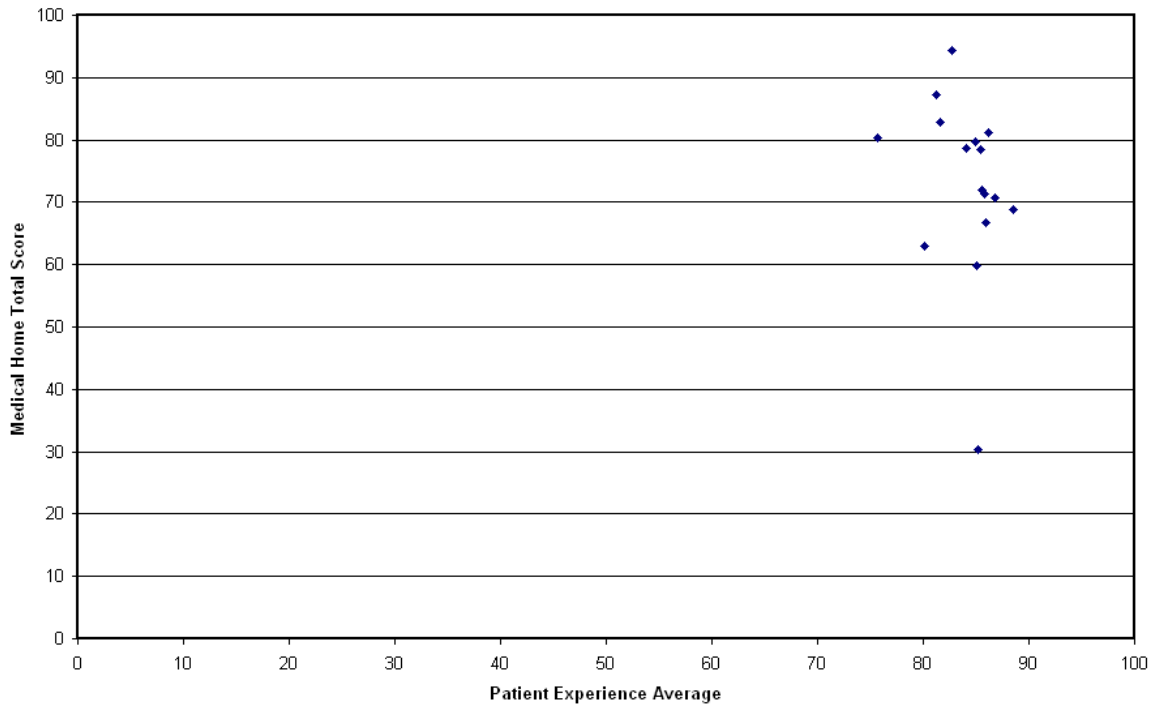

Figure 4.1 Scatter plot of MHQP Average Score and PPCC-PCMH Total Score Per Practice

4.4.2 Hypothesis 2: Organizational Access

The scatter plot of the MHQP organizational access measure (100 maximum) and the PPC-PCMH access and communication elements (9 maximum) for each practice (Figure 4.2) and the calculated Pearson correlation coefficient of -0.72 (Table 4.9) indicate a very strong negative linear association. The hypothesis results indicate that the MHQP Patient Experience scores for organizational access and the PPC-PCMH scores for patient access and communication (Elements 1A and 1B) are negatively correlated with a p-value of 0.002, significant at the 0.05 level. That is, there is a strong significant

negative correlation (r = -0.72) between the MHQP measure of organizational access and the PPC-PCMH elements of Access and Communication. In this study of 16 practices, a correlation of -0.72 was detected with 27% power at the 0.05 significance level.
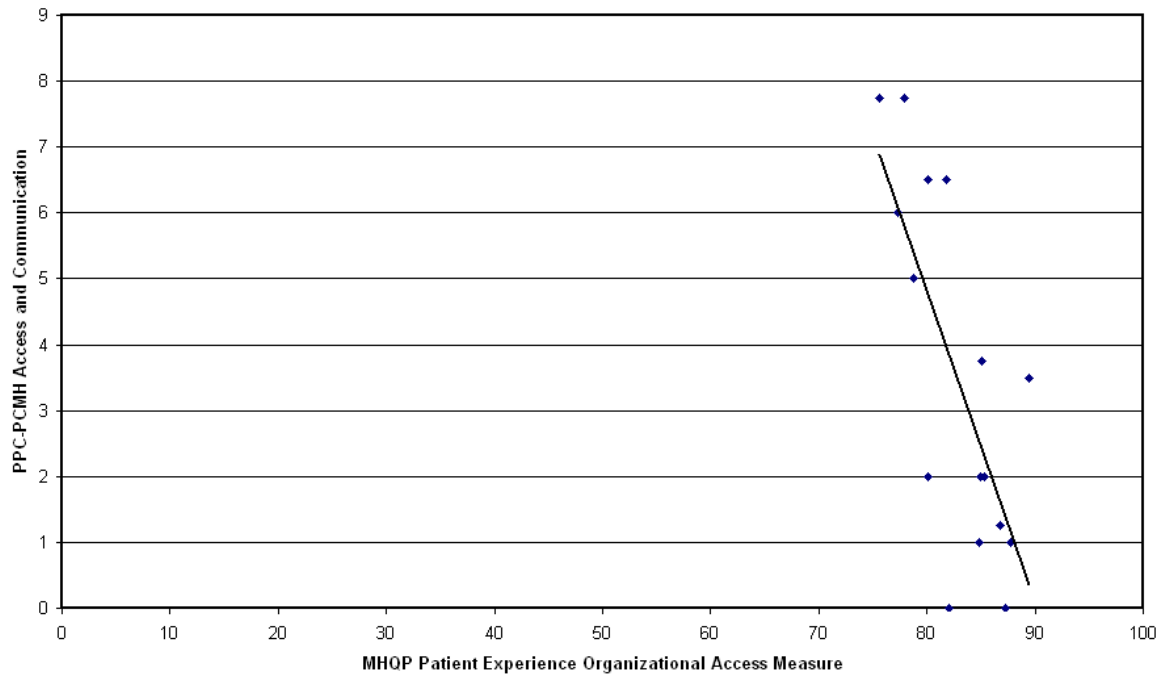

Figure 4.2 Scatter plot of MHQP Organizational Access and PCMH Access and Communication (Elements 1A and 1B)

4.4.3 Hypothesis 3: Health Promotion

The scatter plot of MHQP patient experience health promotion measure (100 maximum) and the PPC-PCMH preventive service clinician reminders scores (4 maximum) for each practice (Figure 4.3) and the calculated Pearson correlation coefficient of 0.17 (Table 4.9) indicate a very weak positive linear association. The hypothesis results indicate that the MHQP patient experience scores for health promotion and the PPC-PCMH scores for preventive service clinician reminders (Element 3B) are positively correlated with a p-value of 0.524, not significant at the 0.05 level. That is, there is no significant correlation between scores. Limited variation in the PPC-PCMH scores (only three values below 4) limited the ability to detect a linear relationship.
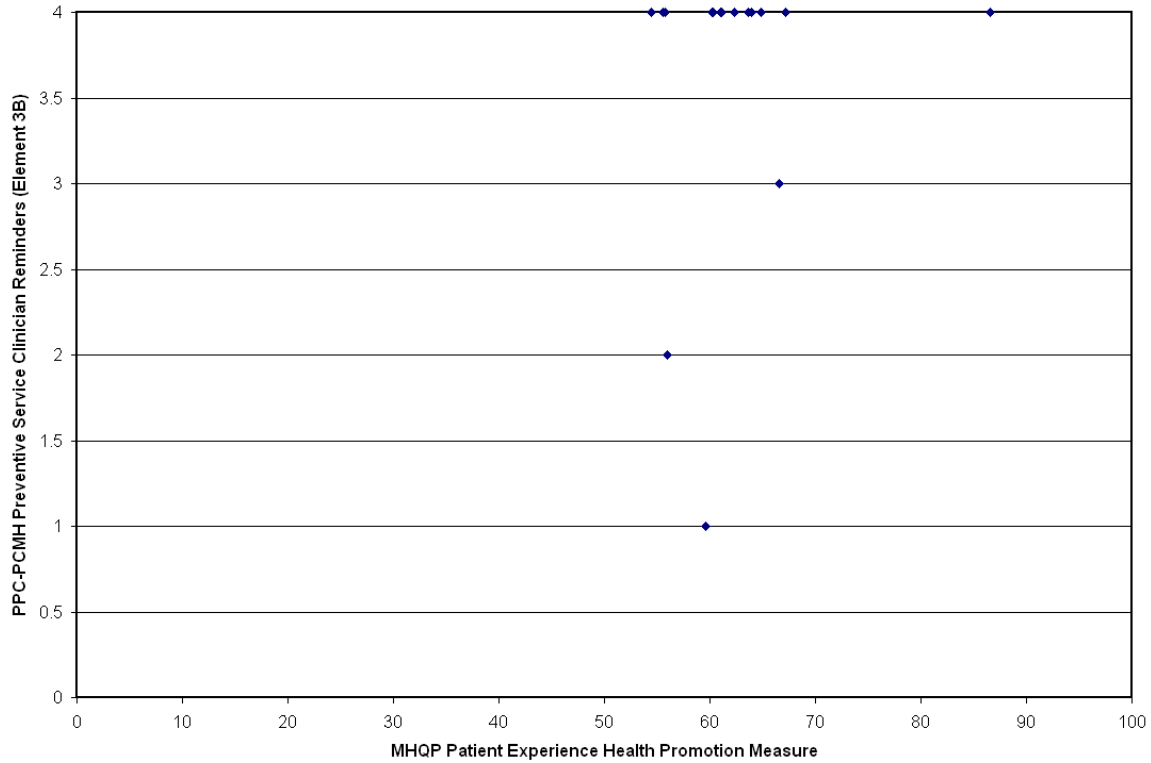
161

Figure 4.3 Scatter plot of MHQP Health Promotion Measure and PCMH Preventive Service Clinician Reminders (Elements 3B)

4.4.4 Hypothesis 4: Clinical Team

The scatter plot of MHQP patient experience clinical team scores (3 maximum) and the PPC-PCMH practice organization scores (100 maximum) for each practice (Figure 4.4) and the calculated Pearson correlation coefficient of -0.49 (Table 4.9) indicate a weak negative linear association. The hypothesis results indicate the MHQP Patient Experience score for clinical team and the PPC-PCMH scores for practice organization (Element 3C) are negatively correlated with a p-value of 0.053, not significant at the 0.05 level.   That is, there is no significant correlation between scores.
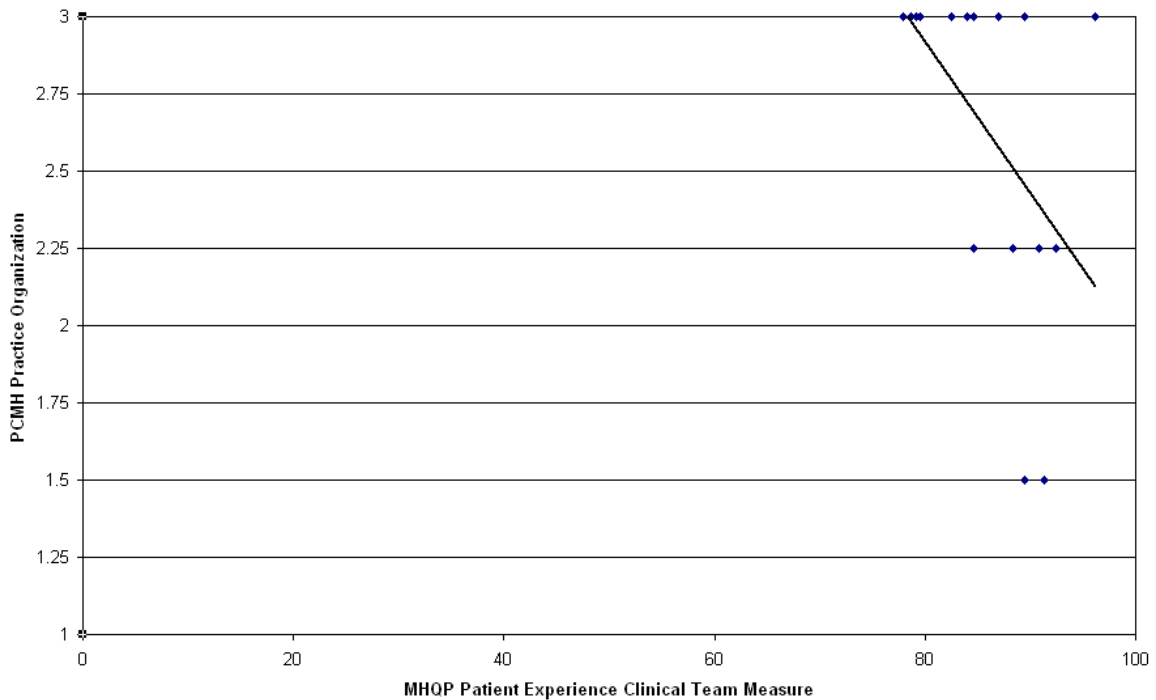
162

Figure 4.4 Scatter plot of MHQP Clinical Team and PCMH Practice Organization
(Elements 3C)

4.4.5 Practice Size as Confounding Variable

Rittenhouse et al. (2008) studied the relationship between Patient-Centered

Medical Home infrastructure elements and medical group size, and found the largest

medical groups have the highest levels of medical home infrastructure. Table 4.10 shows

the final results of the PPC-PCMH survey on level of NCQA recognition by practice

level. As shown, the only three practices scoring at a Level 3 were from large practices.

Table 4.10 PPC-PCMH Recognition Level by Size of Practice (n=16)

	Total Practices	% of Total Practices	Medium Practice, 3-8 PCPs	% of total Medium	Large Practice, >8 PCPs	% of total Large
Level 3	3	19%	0	0%	3	38%
Level 2	0	0%	0	0%	0	0%
Level 1	12	75%	8	100%	4	50%
No Level	1	6%	0	0%	1	13%
Total	16		8		8	

Additionally, Figure 4.5 shows the percent of total possible points combined across practice sizes for the PPC-PCMH elements and total score used in this study. For each element, the percent score is the average raw score across practices divided by the total eligible points for that element. The total score percentage is the average score divided by 100 (maximum score for PPC-PCMH survey). Figure 4.5 suggests that practices with more than 8 primary care physicians (PCP) perform higher on three of the PPC-PCMH elements studied, and most notably on scores indicating the presence of processes and results for access and communication (Element 1A+1B).
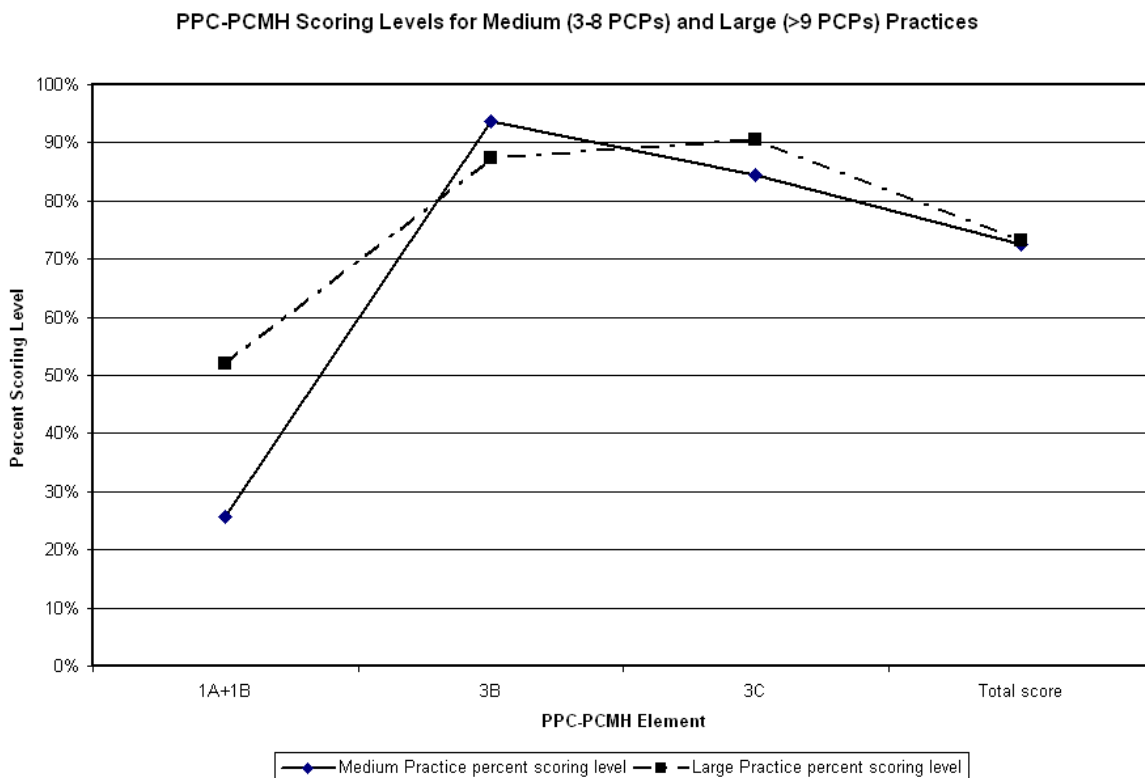

Figure 4.5 PPC-PCMH Scoring Levels for Medium (3-8 PCPs) and Large (>8 PCPs) Practices

We explored practice size as a confounding variable in our analyses by repeating the only significant correlation (hypothesis 2) and stratifying on practice size. The resulting correlations were slightly reduced, to -0.667 for large practices (>8 physicians)

and -0.663 for medium size practices (3-8 physicians).   The scatter plots and Pearson correlation coefficients suggest a negative linear association but hypothesis testing did not find significant associations at the 0.05 level.   The power of the test to reject the null hypothesis that there was no correlation was reduced due to small sample size (8 practices) per study.   Therefore, we cannot say whether the practice size is a significant contributor to the statistically significant negative correlation observed between the MHQP organizational access measure per practice and the PPC-PCMH access and communication elements per practice.   There was not enough power in the test to detect significant correlations at this sample size.

4.5 Discussion

In this study, we investigated correlations between practice scores from patient experience measures conducted by MHQP and structures of PPC-PCMH surveyed by an internal surveyor at one health care organization.   The assessment of Medical Home structures using the PPC-PCMH survey was conducted in late 2008 and early 2009.   The MHQP patient experience survey was conducted in late 2009, creating an appropriate study of structures in place and the corresponding patient satisfaction of such structures.

Reid et al. (2009) found that as the structures of a Medical Home increase, patient satisfaction increases.   The PPC-PCMH Recognition program, sponsored by NCQA, also suggests this relationship and uses levels of recognition to indicate the medical homeness of a practice.   In our study, however, one hypothesized association found to be significant had a negative correlation, contrary to expectation.   That is, as the processes and results increase for patient communication and access, patient experience decreases.   We explore several possible reasons for the negative association.

First, it raises the possibility that the PPC-PCMH survey may not be a useful tool

for assessing the types of processes associated with Patient-Centered Medical Home

principles. For instance, the 12 items in Element 1A (Table 4.4) are equally weighted, so

providing an interactive practice web site (item 10) has the same effect on the element

score as scheduling each patient with a personal clinician for continuity of care (item 1).

While each item contributes to the structures of a Medical Home, the necessity and

magnitude of each item could vary. Berenson et al. (2008) indicate that the PPC-PCMH

tool heavily weights IT technology compared to patient-centered processes like care

coordination. In addition, it is possible that the variation of PPC-PCMH scores do not

differentiate structural characteristics for each element.   Currently, there are no peer

reviewed publications examining the ability of the PPC-PCMH survey to reliably

differentiate such structures.

Another potential cause of the unexpected negative correlation is that the

processes of patient access (Element 1A) and the processes of the care team (Element 3C)

may be newly implemented. The performance of such systems may worsen before getting

better, or patient experience might be driven by earlier experiences. The phenomenon of

worse-before-better is often seen in IT implementations and process redesign (Ross and

Vitale, 2000). Therefore, the sample proportion of patient experience during the

transformation stage might bias the population statistics causing sampling errors across

practices. Even though Nutting et al. (2009) estimated clinics require a three year time

period to transform to a Medical Home, Reid et al. (2009) found an increase in patient

experience after only 12 months of a Medical Home implementation. Staggered

implementation of structures across medical practices prior to the survey can also affect

the interpretation of the results. Carney et al. (2009) reported that many features of a PCMH were already established at the baseline of their study making it difficult to identify major areas for improvement.

Finally, the negative association could represent either poor performance of the processes assessed or other external factors not examined by this study. Newly hired staff, a redesigned process for triaging, or a new call center layout are examples of how the processes might be affected.

Although our result is statistically significant, our study has some limitations. The limited number of practices in our sample also made it difficult to test relationships with statistical significance. Even with the significant finding in access and communication, a sample size of 30 would be needed to achieve a 0.80 power. The sample size also limited the extent to which the effects of multiple covariates could be examined simultaneously, which is why this study emphasizes the use of bivariate and partial correlations through stratification. Larger studies that adjust these associations for multiple confounding variables would be a valuable future study. Additionally, we only studied one organization and not all of the practices, making the findings difficult to generalize to other health care organizations.

With only one internal surveyor of the PPC-PCMH survey, misinterpretation of the internal assessment tool is possible. The probability of this occurrence was minimized by NCQA feedback on the performance of three practices applying for recognition. However, processing errors (e.g., mistakes in mechanical tasks such as arithmetic or data entry) and response errors (e.g., a subject gives an incorrect response and/or surveyor misinterprets) are still possible. Another limitation is the accuracy of the

PPC-PCMH self-audit. A study of the Physician Practice Connections Readiness Survey (precursor to PPC-PCMH) reported that overall agreement with the on-site audit ranged from 40.9% to 96.7% among lead physicians and the accuracy of self-reports of practice systems varied by type of system being assessed and by type of respondent (Scholle et al., 2008).

4.6 Conclusions

This paper provides an exploratory framework for understanding the relationships between the structures assessed by the PPC-PCMH tool and measures of patient experience. With this framework, we have extended the work of Solberg et al. (2008) and Reid et al. (2009) by incorporating the PPC-PCMH tool used to qualify primary care practices as Patient-Centered Medical Homes. This study is important because patient-centered medical homes are fundamentally designed, using seven joint principles created by a collaboration of NCQA and medical societies, to enhance the patient experience, such that positive correlations between structures of a medical home and patient experience should be expected.

However, we found a significant negative correlation at the 0.05 level between the processes and results of access and communication and the MHQP patient experience measure of organizational access. We highlighted several potential causes for this finding including validity of the PPC-PCMH tool, implementation timing of PCMH structures or true performance of assessed structures, and other confounding factors such as process or IT design changes that would negatively influence patient satisfaction. As health care organizations use the PPC-PCMH survey to assess the level of medical

homeness in their practices, considerations of measurement precision and error should be made before interpreting the results.

Other hypothesized correlations did not yield significant results. For instance, we expected to find a significant correlation at the 0.05 level between the MHQP patient measure asking patients how well doctors give preventive care and the PPC-PCMH element indicating a process either paper or electronic existed with guideline-based reminders. In looking at the raw data in Figure 4.3, it is quite possible that a correlation is present but is limited by either sample size or the design of the PPC-PCMH questions lack variation to reliably differentiate practices in health promotion.

While we have presented a framework for analyzing the PPC-PCMH tool, several limitations were highlighted. Sample size was the most significant limitation of this study. The estimated sample size to achieve a power of 0.95 was 40 for the access and communication measure. In addition, the ability to compare across practices can be influenced by variation in self-reported survey data (Scholle et al., 2008) and variation in surveying process. The findings also reflect data from a single health care organization and in particular, use of a convenient sample of practices. Future studies could address these limitations and expand the scope to include other performance measures such as quality and efficiency. It is quite possible that the MHQP patient experience measures are not a good means to evaluate medical home performance in areas of health promotion and clinical team performance.

In this study, we examined whether the constructs of medical homeness, as measured by the PPC-PCMH survey, are associated with patient experience and whether there are other confounding variables such as practice characteristics or measurement

error limiting the ability to detect such correlations. Although the PPC-PCMH instrument is useful in identifying differences in major IT and process structures across practices, much remains to be done to compare the results across practices and to relate them to patient outcomes. This study provides a framework for future studies and supports information-enabled decision making concerning the assessment of a Patient-Centered Medical Home.

4.7 Acknowledgments

4.8 References

Barr, M. S. "The Need to Test the Patient-Centered Medical Home," *JAMA: The Journal of the American Medical Association* (300:7), August 20, 2008, pp. 834-835.

Berenson, R. A., T. Hammons, D. N. Gans, S. Zuckerman, K. Merrell, W. S. Underwood, and A. F. Williams. "A House is Not a Home: Keeping Patients at the Center of Practice Redesign," *Health Affairs (Project Hope)* (27:5), Sep-Oct, 2008, pp. 1219-1230.

Caldis, T. "Composite Health Plan Quality Scales." *Health Care Financing Review* (28:3), 2007, pp. 95.

Carney, P. A., M. P. Eiff, J. W. Saultz, A. B. Douglass, C. J. Tillotson, S. D. Crane, S. M. Jones, and L. A. Green. "Aspects of the Patient-Centered Medical Home Currently in Place: Initial Findings from Preparing the Personal Physician for Practice," *Family Medicine* (41:9), Oct, 2009, pp. 632-639.

Institute of Medicine (U.S.). Committee on Quality of Health Care in America. *Crossing the Quality Chasm : A New Health System for the 21st Century*, National Academy Press, Washington, D.C., 2001.

Kaiser Family Foundation. "Side-by-Side Comparison of Major Health Care Reform Proposals," *Focus on Health Reform,* September 18, 2009. Available from: www.kff.org. Accessed October 20, 2009.

Kohn, L. T., J. Corrigan, and M. S. Donaldson. *To Err is Human : Building a Safer Health System*, National Academy Press, Washington, D.C., 2000.

Landon, B. E., S. L. Normand, E. Meara, Z. Qi, S. R. Simon, R. Frank, and B. J. McNeil. "The Relationship between Medical Practice Characteristics and Quality of Care for Cardiovascular Disease," *Medical Care Research and Review : MCRR* (65:2), Apr, 2008, pp. 167-186.

National Committee for Quality Assurance. "Physician Practice Connections—Patient-Centered Medical Home (PPC-PCMH) Standards," Available at: http://www.cms.hhs.gov/DemoProjectsEvalRpts/. Accessed May, 28, 2009.

Reid, R. J., P. A. Fishman, O. Yu, T. R. Ross, J. T. Tufano, M. P. Soman, and E. B. Larson. "Patient-Centered Medical Home Demonstration: A Prospective, Quasi-Experimental, before and After Evaluation," *The American Journal of Managed Care* (15:9), Sep 1, 2009, pp. e71-87.

Rittenhouse, D. R., L. P. Casalino, R. R. Gillies, S. M. Shortell, and B. Lau. "Measuring the Medical Home Infrastructure in Large Medical Groups." *Health Affairs* (27:5), 2008, pp. 1246.

Rodriguez, H. P., T. von Glahn, W. H. Rogers, H. Chang, G. Fanjiang, and D. G. Safran. "Evaluating Patients' Experiences with Individual Physicians: A Randomized Trial of Mail,

Internet, and Interactive Voice Response Telephone Administration of Surveys," *Medical Care* (44:2), Feb, 2006, pp. 167-174.

Rogers, J. C. "The Patient-Centered Medical Home Movement--Promise and Peril for Family Medicine." *Journal of the American Board of Family Medicine (Formerly Journal of the American Board of Family Practice)* (21:5), 2008, pp. 370.

Rosenthal, M. B., F. S. de Brantes, A. D. Sinaiko, M. Frankel, R. D. Robbins, and S. Young. "Bridges to Excellence--Recognizing High-Quality Care: Analysis of Physician Quality and Resource use." *Am J Manag Care* (14:10), 2008, pp. 670.

Rosenthal, T. C. "The Medical Home: Growing Evidence to Support a New Approach to Primary Care." *Journal of the American Board of Family Medicine (Formerly Journal of the American Board of Family Practice)* (21:5), 2008, pp. 427.

Safran, D. G., M. Karp, K. Coltin, H. Chang, A. Li, J. Ogren, and W. H. Rogers. "Measuring Patients' Experiences with Individual Primary Care Physicians. Results of a Statewide Demonstration Project," *Journal of General Internal Medicine* (21:1), Jan, 2006, pp. 13-21.

Scholle SH, Pawlson LG, Solberg LI, Shih SC, Asche SE, Chou AF, and Thoele MJ. "Measuring Practice Systems for Chronic Illness Care: Accuracy of Self-Reports from Clinical Personnel." *Joint Commission Journal on Quality and Patient Safety* (34:7), Jul-1, 2008, pp. 407-416.

Scholle, S. H., L. G. Pawlson, L. I. Solberg, S. C. Shih, S. E. Asche, A. F. Chou, and M. J. Thoele. "Measuring Practice Systems for Chronic Illness Care: Accuracy of Self-Reports from Clinical Personnel," *Joint Commission Journal on Quality and Patient Safety / Joint Commission Resources* (34:7), Jul, 2008, pp. 407-416.

Schultz, D. V. "The Importance of Primary Care Providers in Integrated Systems," *Hfm (Healthcare Financial Management)* (49:1), 01, 1995, pp. 58.

Senate Finance Committee. "Description of Policy Options: Transforming the Health Care Delivery System: Proposals to Improve Patient Care and Reduce Health Care Costs, April 29, 2009," Available at: http://finance.senate.gov/sitepages/legislation.htm. Accessed May, 28, 2009.

Sequist, T. D., E. C. Schneider, M. Anastario, E. G. Odigie, R. Marshall, W. H. Rogers, and D. G. Safran. "Quality Monitoring of Physicians: Linking Patients' Experiences of Care to Clinical Quality and Outcomes." *Journal of General Internal Medicine* (23:11), 2008, pp. 1784.

Solberg, L. I., S. E. Asche, L. G. Pawlson, S. H. Scholle, and S. C. Shih. "Practice Systems are Associated with High-Quality Care for Diabetes," *The American Journal of Managed Care* (14:2), Feb, 2008, pp. 85-92.

Solberg, L. I., S. E. Asche, S. M. Shortell, R. R. Gillies, N. Taylor, L. G. Pawlson, S. H. Scholle, and M. R. Young. "Is Integration in Large Medical Groups Associated with Quality?" *The American Journal of Managed Care* (15:6), Jun 1, 2009, pp. e34-41.

Starfield, B. "Is US Health really the Best in the World?" *Journal of the American Medical Association* (284:4), July 26, 2000, July 26, 2000, pp. 483-485.

Starfield, B., L. Shi, and J. Macinko. "Contribution of Primary Care to Health Systems and Health," *Milbank Quarterly* (83:3), 09, 2005, pp. 457-502.

Vesely, R. "Medical Home Push. Align Payment with Quality: AHIP." *Modern Healthcare* (38:26), 2008, pp. 15.

Woolliscroft, J. O., J. TenHaken, J. Smith, and J. G. Calhoun. "Medical Students' Clinical Self-Assessments: Comparisons with External Measures of Performance and the Students' Self-Assessments of overall Performance and Effort," *Academic Medicine: Journal of the Association of American Medical Colleges* (68:4), Apr, 1993, pp. 285-294.

Zoffmann, V., I. Harder, and M. Kirkevold. "A Person-Centered Communication and Reflection Model: Sharing Decision-Making in Chronic Care," *Qualitative Health Research,* Jan 25, 2008.

5.0 Summary

Health care improvement is gaining momentum as lawmakers, health care organizations, insurers and payers seek to develop a system with informed, activated patients, and prepared proactive practice teams delivering high quality care at lower costs. This thesis explored three related topics mentioned frequently in the 2010 US health reform proposals – Electronic Health Records (EHR), physician profiling, and Medical Home. The results provide empirical evidence and methodologies to support information-enabled decision making. My dissertation also demonstrates an effective collaboration between engineers and health care professionals to design and implement effective tools to improve health care delivery systems (Lawrence, 2005).

5.1 EHR-enabled Standardization

Using a grounded theory approach, we present a standardization dynamics model (Figure 1.2) that frames the conditions shaping the process state of standardization, encompassing multi-level influencers. This microsystem level model combines the experience of individual providers (Karsh, 2004; Nemeth et al., 2008; Kaplan, 2001) with system models (Berwick, 2002; Carayon et al., 2006; Wagner, 1998) to support organizational decision making regarding IT and process design.

Four pathways (Figure 1.3) surfaced in the data as intuitive processes initiated by physicians to deliver care across different levels of patient complexity and routine or complex information exchanges. The value of such pathways is that they frame choices about how to use embedded EHR structures to support effective patient care plans, going beyond generic advice to redesign processes in conjunction with IT implementation. A process design tool was created to qualitatively describe the process state of standardization for current state and future state analysis.

There are some limitations to our research. The data collected was from one health care organization using one EHR system. Additional longitudinal research should test the viability of the standardization dynamics model, as well as the interactions and relationships between conditions and consequences along pathways, for different EHR platforms and for different organizational and societal contexts. Additionally, our discussion on designing for pathways assumes that individual style is easily influenced. Future work could explore the possibility of expanding the four pathways to include properties of individual style. Our study contributes to a better understanding of the process state of standardization and how medical management should think about, and make decisions about, EHR embedded structures and process design in primary care.

5.2 Physician Profiling

For this study, we designed a process for creating a physician profiling model, supporting key design criteria suggested in the literature. The design criteria are validity, purpose, completeness and accuracy of the data, reliability, and usefulness.
A physician profiling model was developed at a case study site called Health Clinic using the design process, with an emphasis on individual reliability estimates. Twenty-four measures using claims data from multiple payers were combined into domains of quality of care and efficiency. The measures were combined in each domain using standardized Z scores, which addressed the differences in measure difficulty and the differences in patient panel characteristics by specialty. The results showed that reliable physician profiling is possible across care domains using a hierarchical composite model. The model was used for an incentive program for payout across 199 primary care physicians

at Health Clinic, demonstrating its effectiveness and providing a basis for use in other health care contexts.

Limitations of the model building process include long delays in feedback due to using claims data, an inability to recognize performance improvement over previous years, and physician feedback actionable at the domain level rather than the individual measure level. The model developed at the case study site was limited by the inclusion of only some payers, an inability to bridge the organization rate with the average physician rate, and the exclusion of physicians with a panel size less than 30. Broadly, physician profiling requires initiatives to create shared pools of transparent physician performance information from all payers and systems, to broaden the eligible patient population per measure and per physician.

A future study could investigate the variation of reliability estimates, the misclassification of physicians as a function of reliability, and methodologies for the disposition of scores with individual reliabilities less than a desired threshold. Another future study could investigate the limitations of using weighted Z scores across heterogeneous populations by examining varying minimum quality event thresholds of individual measures. Other applications of the physician profiling model can be explored. For instance, the use of case-mix adjustments for across patient population studies could allow for accurate physician profiling for applications. Such studies will help advance the usefulness and practical knowledge of physician performance profiling in a health care organization, health plan or federal programs.

5.3 Patient-Centered Medical Home

This study provided an exploratory framework to understand relationships between the structures assessed by the PPC-PCMH tool and measures of patient experience. The results showed a significant negative correlation for one hypothesized relationship, patient communication and access. Reasons for the negative correlation might include: the validity of the PPC-PCMH tool, implementation timing of PCMH structures on true performance of assessed structures, and other confounding factors such as process or IT design changes that would negatively influence patient satisfaction.

Sample size was a major limitation in this study. The power of the patient communication and access test was only 0.28, or a 28% probability of rejecting the null hypothesis when the null hypothesis is false. For a power of 0.80, a sample size of 30 was needed. With such a low power of the test, even large correlations can go undetected. Another limitation of this study was associated with data collection. The data was drawn from one organization and did not include all practices. This convenience sampling technique can produce inaccurate conclusions about the entire organization and can limit the ability to generalize to other organizations. Also, one internal surveyor was used at the health care organization, which supports consistency across surveys but might produce a bias in the results through misinterpretation of the survey details. Scholle et al. (2008) also reported that the ability to compare across practices can be influenced by variation in self-reported survey data.

Opportunities for future studies include expanding the participating sites at Health Clinic or collaborating to gather data across health care organizations. If controls are put in place for consistency in administering the PPC-PCMH survey, then other clinical sites

within Massachusetts could collectively study relationships between the survey results

and MHQP patient experience data. Additionally, future studies with larger sample sizes

could study the confounding effects of practice characteristics such as size or patient

population on the results of PPC-PCMH surveys and the corresponding changes to the

relationships. Other future studies could examine relationships between the PPC-PCMH

model and other performance measures such as quality and efficiency. This study

supports information-enabled decision making concerning the efficacy of a Patient-

Centered Medical Home.

5.4 Presentations and Awards

This work has been presented at various seminars and conferences and recognized

by different organizations. A summary of the significant presentations, posters and

awards is provided in the following sections.

5.4.1 Presentations

Pelletier, L., Johnson, S., Westrick, E., Klugman, R., Fontaine, E., Krinsky, A., Physician Performance Modeling in a Healthcare Organization. *WPI Innovation Presentation Competition*, Worcester, MA, April 1, 2010.

Pelletier, L., Johnson, S., Westrick, E., Klugman, R., Fontaine, E., Krinsky, A., Zhou, Q., Physician Performance Modeling in a Healthcare Organization. *INFORMS Annual Meeting*, San Diego, CA, October 11, 2009.

Strong, D., Tulu B., **Pelletier L.,** Johnson, S., Bar-On, I., Volkoff O., Garber L., Kashyap, N., Trudel J., EHR Affordances and their Actualizations. Extended Abstract, *SigHealth Workshop, Americas Conference on Information Systems*, San Francisco, CA, August 6, 2009.

Strong, D., Volkoff, O., Johnson, S., Bar-On, I., **Pelletier, L.**, Electronic Health Records and the Changing Roles of Health Care Professionals: A Social Informatics Perspective. *15th Americas Conference on Information Systems*, San Francisco, CA, August 6-9, 2009.

Johnson, S., **Pelletier, L.**, Strong, D., Bar-On, I., Trudel, J., Kashyap, N., Garber, L., Standardization of Health Care Delivery Processes: A Qualitative Study of the Effects of Electronic Health Record Systems. Extended Abstract, *INFORMS MSOM Conference and Special Interest Group (SIGs) Meetings*, MIT, Cambridge, MA, June 29-30, 2009.

Pelletier, L., Johnson, S., Westrick, E., McBride, J., Klugman, R., Fontaine, E., Krinsky, A., Cernak, C., Medical Home and Patient Satisfaction, Why Aren't They Correlated? *UMass Memorial Health Care: Quality and Patient Safety Grand Rounds*, Worcester, MA, June 17, 2009.

Pelletier, L., Johnson, S., Westrick, E., Klugman, R., Fontaine, E., Krinsky, A., Zhou, Q., Kim, R., Assessing Physician Quality and Efficiency in the Ambulatory Setting, a New Model. *UMass Memorial Health Care: Quality and Patient Safety Grand Rounds*, Worcester, MA, June 17, 2009.

Pelletier, L., Westrick, E., Johnson, S., McBride, J., Fontaine, E., Examining the Relationship between the Medical Home and Quality Outcomes. *Institute of Industrial Engineers Annual Conference and Expo*, Miami, May 2009.

5.4.2 Posters

Pelletier, L., Johnson, S., Westrick, E., Klugman, R., Fontaine, E., Krinsky, A., Physician Performance Modeling in a Healthcare Organization. *Worcester Polytechnic Institute Graduate Research Achievement Day Poster Symposium*, Worcester, MA, March 31, 2010.

Pelletier, L., Johnson, S., Strong, D., Tulu, B., Bar-On, I., Trudel, J., Kashyap, N., Garber, L., Standardization of Health Care Delivery Processes: A Qualitative Study of the Effects of Electronic Health Record Systems. Poster AMIA-0397-A2009, *American Medical Informatics Association 2009 Annual Symposium*, San Francisco, November 14-18, 2009.

Pelletier, L., Johnson, S., Westrick, E., Krinsky, A., Klugman, R., Fontaine, E., Zhou, Q., Kim, R., Physician Performance Modeling in a Healthcare Organization. *Institute of Industrial Engineers Annual Conference and Expo*, Miami, May 2009.

Pelletier, L., Johnson, S., Strong, D., Bar-On, I, Standardization of Health Care Delivery Processes: A Qualitative Study of the Effects of Electronic Health Record Systems, *Worcester Polytechnic Institute Graduate Research Achievement Day Poster Symposium*, Worcester, MA, March 28, 2008 and April 1, 2009.

5.4.3 Awards

- WPI Innovation Presentation Competition: Judge's Special Recognition for Project with the Largest Societal Impact (2010)

- WPI Graduate Poster Competition: 3rd place in Science Division (2010)

- IIE Annual Conference and Expo: Doctoral Colloquium participant (2009)

- AMIA Annual Symposium: Doctoral Colloquium participant (2008)