Examination of the Use of Electronic Health Record Data for Measuring Performance in Diabetes Care

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in the Graduate School of The Ohio State University

By

Annemarie Gregory Hirsch, M.P.H., B.A.

Graduate Program in Public Health

The Ohio State University

2012

Dissertation Committee

Professor Amy K. Ferketich
Professor Ann S. McAlearney
Professor Eric E. Seiber
Professor Philip F. Binkley, Advisor
Abstract

There is enthusiasm among policymakers for the potential for electronic health record (EHR) systems to drive improvements in health care quality. The federal government has invested billions of dollars in an effort to expand EHR technology. This investment goes beyond the adoption of an EHR system. The government is providing incentives for healthcare providers to leverage EHR data for performance measurement. However, EHR data and this new application of EHR data are poorly understood. This study was designed to gain an understanding of EHR data and the validity of using it to measure performance in diabetes care.

This was a two-part project. The first part consisted of interviews with 17 clinician users of an EHR system to learn how data is entered and processed in the EHR and to determine what factors influence data entry practices. During these interviews we collected data on how data entry practices impact performance measurement in diabetes care. The second part of this study analyzed data extracted from an EHR to assess the validity of eight different methods of identifying the target population for diabetes performance measures (i.e. patients with diabetes). The analysis assessed each of the methods for sensitivity and specificity, using a manual chart review as the gold standard. Next, we calculated performance on three diabetes measures (cholesterol, blood pressure, glucose control) to determine whether the method of identifying the target population was associated with the outcome of these performance measures.
Clinicians endorsed the use of the problem list in the EHR to identify the target population for diabetes measures. Clinicians believed that organizational factors, such as pressure from leadership, workflow, and quality improvement programs ensured the accuracy of the problem list, while making the encounter diagnosis more likely to over- or under-identify patients with diabetes. The clinicians also noted the unintended consequences of using EHR data for performance measurement and quality initiatives. They felt this application of EHR data could result in improved documentation in the EHR, insurance coverage issues, unnecessary patient anxiety, and pressure among physicians to apply quality standards to patients to whom such standards should not apply.

All 8 of the EHR-based methods for identifying patients with diabetes had high specificity (>99.5%) and moderate to high sensitivity (65% to 100%). The method of identifying patients with diabetes did not appear to have a large impact on the outcomes of the performance measures. However, the EHR criteria used in each of the definitions selectively identified patients who had better quality performance scores. Specifically, definitions that classified patients as having diabetes if they had at least two diabetes encounter diagnoses selected patients with better performance and excluded patients with poorer performance.

The findings from this study provide important information to administrators of quality reporting programs; health system leaders participating in performance measurement; and patients, providers and payers viewing performance data. These findings have implications beyond performance measurement. Health services researchers and epidemiologists can benefit from a better understanding of EHR data and its use in identifying disease populations.
Dedication

For Andy and Rory
Acknowledgements

I owe the success of this project to a number of people and organizations. First, I am grateful to the Agency for Healthcare Quality and Research for funding this project through its Health Services Research Dissertation Award, grant 1R36 HS020165-01. This funding allowed me to fully dedicate myself to this project over the past year.

I would like to thank Dr. Walter F. Stewart, Director of the Center for Health Research at Geisinger. For the past seven years, Dr. Stewart has provided me mentorship in the fields of epidemiology and health information technology (HIT) research. He introduced me to the area of HIT research and was an excellent sounding board for ideas as I developed my proposal for this project. He helped facilitate my access to Geisinger’s electronic health record data and provided key connections to individuals in the quality office who assisted me with subject recruitment for the clinician interviews. Dr. Stewart’s support of this project has been critical to its success.

Dr. Ann McAlearney has been another important mentor to me in the field of HIT and taught me everything I know about qualitative research methods. She gave me the opportunity to participate in a number of projects and to contribute to research publications in the field of HIT. She took a generous amount of time to introduce me to new research topics and research methods that I used for this project. Throughout my doctoral education, Dr. McAlearney has provided instruction and advice that has been essential to my completion of the doctoral program. I am very grateful for her mentorship.
I am also very thankful to Dr. Amy Ferketich who has provided me with tremendous guidance and support since I arrived at The Ohio State University. Over the years I have turned to Dr. Ferketich countless times for advice on classes, my dissertation, and my career. She has always provided encouragement and taken the time to ensure my success in the program. Dr. Ferketich has been there for me at each phase of my doctoral program, most recently providing frequent, timely, and immensely helpful feedback on this final project.

My other committee members, Dr. Philip Binkley and Dr. Eric Seiber have also played very important roles in the development of this project. Dr. Binkley’s willingness to let me pick his brain regarding his own experience with the electronic health record as a physician was an inspiration for the physician interviews I conducted for this project. Dr. Seiber’s thoughtful comments, both during the proposal phase of this project and on earlier drafts of this paper, helped to broaden my thinking on this area of research.

Finally, I would never have been able to complete this project without the support of family and friends. I am particularly thankful for my parents, who knew I would go back and get my PhD before I did. I am grateful to my husband, Andy, who has been a constant source of encouragement throughout my doctoral work. Finally, I am thankful for my son, Rory, who provided a very welcome distraction to the writing process.
Vita

March 24, 1978 ............................ Born – Abington, PA

2000 .............................................. B.A. Biology and Psychology,  
La Salle University

2000 – 2001 .................................... Research Assistant,  
Center for Applied Research in Aging and Health  
Thomas Jefferson University

2001 – 2002 .................................... Research Assistant,  
Center for the Psychosocial Research of Health  
and Illness  
Columbia University

2003 .............................................. Master of Public Health  
Columbia University

2003 – 2004 .................................... Research Coordinator  
Institute on Aging, Alzheimer’s disease Center  
University of Pennsylvania

2005 – 2012 .................................... Project Manager III  
Center for Health Research  
Geisinger Health System

2008 .............................................. University Fellow  
The Ohio State University

2009 – 2010 .................................... Graduate Teaching Assistant  
Department of Epidemiology  
College of Public Health  
The Ohio State University

2010 – 2011 .................................... Graduate Research Assistant  
Department of Epidemiology  
College of Public Health  
The Ohio State University


Fields of Study

Major Field: Public Health

Specialization: Health Services Management and Policy
Table of Contents

Abstract .............................................................................................................................. ii
Dedication ........................................................................................................................ iv
Acknowledgements .......................................................................................................... v
Vita...................................................................................................................................... vii
List of Tables..................................................................................................................... xii

Chapters:

1. The Electronic Health Record: A New Source of Data for Performance Measurement ............... 1
   1.1 The Electronic Health Record: Beyond Clinical Care ................................................. 2
   1.2 Performance Measurement for Quality Reporting ..................................................... 5
   1.3 Performance Measurement: From Claims to EHR Data .............................................. 6
   1.4 EHR Concerns ........................................................................................................... 8
   1.5 Measuring Performance in Diabetes Care ................................................................. 10
   1.6 Description of Project .............................................................................................. 14

2. Organizational Influences on Data Entry Practices in the Electronic Health Record: Implications for Performance Measurement..... 16
   2.1 Background ................................................................................................................ 16
   2.2 Purpose ....................................................................................................................... 17
   2.3 Methods ..................................................................................................................... 17
      2.3.1 Research Design ................................................................................................. 18
      2.3.2 Setting and Recruitment ..................................................................................... 18
      2.3.3 Study Participants .............................................................................................. 18
      2.3.4 Measures ............................................................................................................ 19
      2.3.5 Analysis .............................................................................................................. 20
   2.4 Results ...................................................................................................................... 20
      2.4.1 Documenting Diabetes Diagnoses ........................................................................ 21
      2.4.2 Clinician Recommendations for EHR Fields to Use in Identifying Patients with Diabetes ................................................................. 22
      2.4.3 Organizational Factors Influencing EHR Data Entry ............................................. 24
         2.4.3.1 Workflow ...................................................................................................... 24
         2.4.3.2 Internal Quality Improvement Program ........................................................... 27
List of Tables

Table 1.1 EHR Patient Record Data Categories .................................................. 3

Table 1.2 PQRI 2010 EHR Specifications (Abridged) for Diabetes
Quality Measures ................................................................................................. 12

Table 1.3 EHR-Based Definitions for Diabetes: Sensitivity and Specificity .......... 13

Table 2.1 Clinician Comments Regarding Documentation of Diabetes
Diagnoses in the Problem List and Encounter Diagnoses Fields ................. 24

Table 2.2 Clinician Comments Regarding Organizational Influences on
Problem List and Encounter Diagnoses ............................................................ 26

Table 3.1 EHR-based Definitions for Identifying Primary Care Patients
with Diabetes ........................................................................................................ 37

Table 3.2 Sensitivity and Specificity of Definitions for Diabetes ..................... 40

Table 3.3 Proportion of Patients Identified by Diabetes Definitions Meeting
CMS Quality Standards: Patients Identified by Definitions and
Patients Missed by Definitions ........................................................................ 42

Table 4.1 Comments Regarding Negative Unintended Consequences
of Using an EHR Algorithm to Identify Diabetes Patients
for Quality Measures ......................................................................................... 58
The United States health system is suffering from a quality crisis.\textsuperscript{1} Despite having the most costly health system in the world, the U.S. underperforms on most dimensions of performance, relative to other countries.\textsuperscript{2} Over the past few years, policymakers have hailed the electronic health record (EHR) as the answer to the health care quality crisis in America. The President of the United States, Barack Obama, said the EHR will “improve the quality of our health care while lowering its cost” and claimed it will “save lives by reducing the deadly but preventable errors that pervade our health system.”\textsuperscript{3} However, simply increasing EHR adoption will be inadequate to substantially improve care.\textsuperscript{4} EHR data must be leveraged for secondary purposes, such as performance measurement, in order to make strides in quality improvement. The focus of this paper is on the use of EHR data for measuring performance for quality reporting.

Policy makers emphasize that valid performance measurement is essential to the success of health reform.\textsuperscript{5} Performance measures drive quality initiatives such as public reporting of physician performance and pay-for-performance (P4P) programs.\textsuperscript{6,7} Performance measures help identify gaps in healthcare delivery and target areas for improvement. Invalid performance measures could undermine potentially effective quality initiatives and misinform providers and hospital leaders regarding the quality of care they are providing. As a result, health care providers may become overconfident about the quality of care provided and reduce improvement efforts, introducing harm.\textsuperscript{6}
For more than a decade, organizations such as The Centers for Medicare and Medicaid Services (CMS) have used claims data to derive performance measures for their quality initiatives. While medical records were only available in paper form, accessing clinical data from the medical record was too costly and time intensive to use. The EHR provides a new, cheaper, faster, and easier way to obtain detailed clinical data. As a result, organizations are using EHR data in addition to claims data for their performance measures. Unlike claims data, the use of EHR data for this purpose is relatively new. As a result, we know very little about valid methods of leveraging EHR data for measuring performance.

This document describes a project designed to assess the validity of using EHR data for performance measurement. Diabetes quality measures were used as examples throughout the project, as diabetes quality measures were among the first measures transitioned to EHR-based reporting. The project aims to gain deeper knowledge of EHR data and how it is entered in order to understand how data entry practices can impact performance measures. The project also evaluates multiple methods of using EHR data to measure performance in diabetes care. The next section provides background information on EHR adoption, the history of performance measurement, and the potential challenges of using EHR data in performance measurement for quality reporting.

1.1 The Electronic Health Record: Beyond Clinical Care

An electronic health record (EHR) is a longitudinal electronic record of patient health information generated by clinical encounters in a care delivery setting. Data elements can include demographic data, diagnoses, medications, clinical orders, past medical history, immunizations, and laboratory results, as detailed in Table 1.1. The use of EHR data elements can go beyond the clinical encounter. Data can be used for such
important secondary purposes as comparative effectiveness research, clinical decision support, and performance measurement. This document focuses on the use of EHR data for measuring performance for quality reporting.

Table 1.1 EHR Patient Record Data Categories

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td>Personal data elements, sufficient to identify the patient, collected from the patient, and not related to health status or services provided.</td>
</tr>
<tr>
<td>Legal Agreements</td>
<td>Data elements indicating legally binding directions or restraints on patient healthcare.</td>
</tr>
<tr>
<td>Financial Information</td>
<td>References to the financial bodies that will cover the cost of the care.</td>
</tr>
<tr>
<td>Provider/practitioner</td>
<td>Contains the descriptive data about each provider and may then be referenced when recording data about the events of health care.</td>
</tr>
<tr>
<td>Problem List</td>
<td>Constitutes a master list of all a patient’s problems or diagnoses. This list includes clinical problems, a diagnosis summary and stressor exposure, an ongoing list of clinically significant health status events, resolved and unresolved, in a patient’s life.</td>
</tr>
<tr>
<td>Immunization</td>
<td>A component of health history that contains, chronologically, all immunizations administered to the patient and their current status.</td>
</tr>
<tr>
<td>Hazardous stressor exposure</td>
<td>The what, where, when, and how data on actual or potential exposure to all biological, physical, or chemical agents that might be associated with adverse health effects.</td>
</tr>
<tr>
<td>Health history</td>
<td>The long-term relevant natural family and patient history and signs which would aid practitioners in predicting or diagnosing illness.</td>
</tr>
<tr>
<td>Assessments</td>
<td>The observations of the provider during structured and systematic examinations of the patient’s body during encounters.</td>
</tr>
<tr>
<td>Clinical orders</td>
<td>Data entries that direct a patient’s treatment, including details on the deliverance of orders and compliance with any diagnostic or therapeutic treatment plans.</td>
</tr>
<tr>
<td>Diagnostic tests</td>
<td>Documentation of the results from the clinical laboratory, radiology, nuclear medicine, pulmonary function and any other diagnostic test.</td>
</tr>
<tr>
<td>Medications</td>
<td>A list of all long-term medications and significant details on all medications prescribed and/or administered in the course of an encounter.</td>
</tr>
<tr>
<td>Scheduled appointments/events</td>
<td>List of planned or scheduled appointments.</td>
</tr>
<tr>
<td>Encounters</td>
<td>A patient session with a practitioner during which information about the patient’s health status is exchanged. Encounter diagnoses consist of a list of all conditions co-existing at the time of the encounter that affect the treatment received or length of stay. A condition of sufficient significance to warrant inclusion for investigative medical studies.</td>
</tr>
</tbody>
</table>

A study commissioned by the Department of Health and Human Services Office of National Coordinator for Health Information Technology (ONC) estimated that less than ten percent of U.S. hospitals have, at minimum, a basic EHR system. Among hospitals who have implemented an EHR, there is great variety as to the level of implementation, ranging from fewer than three ancillary services (i.e., laboratory, radiology, pharmacy) installed to full implementation of EHR-based clinical decision support. To increase the lower than expected EHR adoption rates, the federal government established Medicare incentive payments for “meaningful use” of certified EHR technology. In 2009, the ONC announced nearly $2 billion worth of new programs to help health care providers become meaningful users of electronic health records (EHRs). Eligible health care professionals can generally earn up to $44,000 in extra payments between 2011 and 2015 if they become meaningful users.

To be considered a “meaningful user” of the EHR, providers must do more than implement or use an EHR system. Meaningful use consists of three components. The provider must use an EHR in a meaningful manner, such as e-prescribing; the provider must use an EHR for electronic exchange of health information to improve health care; the provider must use an EHR to submit clinical quality measures.

Federal incentives for EHR adoption, accompanied by the meaningful use requirement to use an EHR for quality reporting, have triggered the rapid adoption of EHR-based performance measures for quality reporting. Administrators of quality programs, such as CMS, have started to accept EHR data for their quality metrics. The next sections briefly describe performance measurement and the movement towards using EHR data.
1.2 Performance Measurement for Quality Reporting

In the late 1990s the Institute on Medicine (IOM) released a report highlighting the “serious and widespread” quality problems that exist in the United States Health Care system. Since the publication of this report the IOM has consistently called for a comprehensive response to improve the quality of health care and has identified performance measurement as a critical step towards transforming health care quality. Performance measures include measures of health care process (e.g., periodic blood and urine tests for diabetic patients), patient outcomes (e.g., 60-day survival rate for cardiac bypass patients), patients’ perceptions of care (e.g., experience with patient-provider communication), and organizational structure and systems associated with the ability to provide high-quality care (e.g., medication order entry systems).

Government agencies, such as the CMS, the Agency for Healthcare Research and Quality (AHRQ), as well as private sector groups, such as the National Quality Forum (NQF), have generated hundreds of performance measures over the last decade. These performance measures support a number of quality improvement initiatives, including public reporting and pay-for-performance (P4P) programs. In public reporting programs, information about the performance of hospitals, health professionals, and health care organizations is released into the public domain in the form of report cards, provider profiles, or consumer reports. The goal of these types of programs is to promote accountability and aid consumers in making informed choices. P4P programs use incentives to encourage providers to adopt evidence-based health care delivery processes. While there is some question as to the effectiveness of these quality initiatives, the IOM reports that these types of programs offer significant promise and are essential to the transformation of the healthcare payment system.
There is consensus that the success of these quality programs is dependent, in part, on the accuracy of the performance measures used to drive them. Inaccurate measures can seriously undermine the credibility of these programs. Invalid measures in P4P programs can result in over or underestimating performance. As a result, payers could unfairly penalize or reward providers and hospitals. Public reporting programs that use invalid performance measures can drive patients to choose care according to misinformation and to make poor care decisions. Finally, a lack of valid measures makes it impossible for providers and hospitals to identify areas of care delivery on which they need to improve, resulting in the perpetuation of poor health care delivery practices.

In designing a valid performance measure, there are a number of important considerations. Among those considerations is deciding what data source to use for assessing performance. The next sections provide a brief history of the sources of performance measure data and the recent transition to incorporating EHR data into performance assessment.

1.3 Performance Measurement: From Claims to EHR Data

Historically, claims data have primarily been the source of data for performance measures. Claims data files are created by payers from bills generated by providers seeking payment for services rendered. Sources of claims data include private insurers, Medicare, Medicaid, the Department of Defense, and the Department of Veterans Affairs. Generally, claims data include demographic data and limited clinical data, including diagnosis and procedure codes as well as corresponding dates. Claims data have been an appealing source of data because data are readily available, are inexpensive to acquire, are computer readable, and contain information about large populations over a period of time.
The literature also reveals important drawbacks to using claims data for performance measurement. For example, claims lack important clinical data. Most administrative files aim to minimize data collection and contain the minimum amount of information required to perform the relevant administrative function (e.g., verify and pay claims).\textsuperscript{10} Claims databases are also subject to missing data issues, because not all services received are billed, patients change insurance payers, a high percentage of US patients do not have stable insurance coverage, and some services do not incur claims.\textsuperscript{30} In addition, disease prevalence can often be underestimated using claims data because the documentation is likely to focus on the most important diagnoses pertaining to the use of health resources.\textsuperscript{9,29} Finally, the accuracy of claims data has been highly debated. The literature reports that agreement rates between claims and clinical data range from 9 to 87\%.\textsuperscript{29,31}

Despite these shortcomings, until recently, claims data were the preferred source for accessing data to measure performance. The cost of extracting clinical data from paper medical records was just too high.\textsuperscript{10} With the federal government’s recent push to increase adoption of EHR systems, however, more organizations are turning to EHR data as a source of data for performance measurement. For example, the National Committee for Quality Assurance (NCQA) is working closely with the ONC, CMS, and NQF to translate Healthcare Effectiveness Data and Information Set (HEDIS) measures into EHR-based measures. Historically HEDIS measures, measures used to monitor performance in more than 90 percent of America’s health plans, have depended on claims data.\textsuperscript{32} Similarly, CMS is transitioning from claims to EHR-based reporting for its pay-for-performance program, The Physician Quality Reporting Initiative (PQRI).\textsuperscript{8,33,34} In the past two years the number of PQRI measures that accept EHR data has grown from 10 to 51.\textsuperscript{33,34}
The biggest advantage of EHR data over claims data is the availability of detailed clinical data. Medical records contain clinical data elements that are not contained in claims data sets, including lab results and vital signs. The availability of more clinically relevant data in an electronic queryable format represents a new source of data that can be leveraged without the expense of a manual chart review. However, as a new data source, methods of using EHR data for performance measures have not been fully vetted or validated. Section 1.4 describes potential challenges to using this new data source for performance measurement in quality reporting.

1.4 EHR Concerns

While the range EHR data and the availability of detailed clinical information make it an attractive data source for quality reporting, there are important issues to consider when using EHR data for quality reporting. This paper focuses on three of these issues. First, the accuracy of EHR data is dependent on the providers who enter the data. Second, there is a range of data elements from which to choose to identify the target population for the quality reporting. Third, as with any new application of health information technology (HIT), there is the potential for negative unintended consequences. These issues are discussed in more detail below.

The accuracy of EHR data is dependent on the health care providers entering data into the medical record. How providers enter data in the EHR can be influenced by factors such as workflow, billing processes, and reimbursement policies and user-interface of the EHR system. Even when there are uniform and discrete data fields available within the same EHR system, physicians tend to use the diverse features of an EHR system in many different ways. As a result, clinical information is recorded in different physical locations of the record (e.g., problem list, progress notes). In some cases, physicians working in practices with EHR capabilities are not utilizing the EHR at...
all.\textsuperscript{46} When choosing what aspects of the EHR can be used for quality reporting, one needs to consider the variation in how physicians use the record and what motivates their documentation patterns. To date, there is little information regarding whether EHR data entry and data processing influence performance measures for quality reporting.

As described in table 1.1, the EHR has a range of data elements available and many of these elements can be used to identify a target population for a quality measure, that is to identify the population of patients whose quality you are trying to measure.\textsuperscript{27} The method for identifying the target population should have high sensitivity and specificity. Misclassifying patients as not having a disease could result in missed opportunities for quality improvement, while misclassifying patients as having disease could result in unnecessary clinical tests and interventions. In addition, the method for defining the target population should not influence the outcome of the performance assessment. In other words, the method should not selectively identify patients who are more or less likely to meet performance goals, as this could result in under- or overestimating performance.\textsuperscript{27}

Given the abundance of data in the EHR, it is not always clear what data elements should be used to identify a target population. Diagnoses are recorded in a number of sections, including the problem list, encounter diagnoses, and progress note. In some cases, medications indicated for only one disease could be used to identify patients with that particular disease. In addition, lab values can be used to identify a disease population for quality measurement. The validity of using these data elements, or a combination of these data elements, to identify a target population is unknown. There are no standards for identifying disease populations using EHR data. In moving to EHR-based measures, CMS has developed specifications for identifying disease populations; however, CMS provides little evidence of the validity of these specifications.\textsuperscript{33}
Finally, as the use of EHR data for quality measures is a relatively new application of HIT, it is important to assess the potential for unintended consequences. There is extensive literature describing the unintended consequences of a range of HIT applications.\textsuperscript{47-54} The negative consequences of this technology range from inefficient workflows to the extreme of patient death.\textsuperscript{47,49,50,51,52} Illustrating the potential seriousness of HIT unintended consequences, Weiner and colleagues coined the term “e-iatrogenesis” defined as patient harm caused at least in part by the application of HIT.\textsuperscript{51} There is little information in the literature about the unintended consequences or the potential for e-iatrogenesis when using the EHR for performance measures in quality reporting.

1.5 Measuring Performance in Diabetes Care

Diabetes is a condition frequently included in the performance assessment and performance improvement activities of organizations such as the NCQA and NQF. A recent review of diabetes performance measures identified 146 distinct measures spanning 31 clinical processes or outcomes.\textsuperscript{55} This attention towards measuring performance in diabetes care is primarily driven by three factors: 1) the importance of the condition, 2) the potential for quality improvement, and 3) the degree to which health care professionals can control mechanisms for improving care.\textsuperscript{56} These are also the reasons we selected performance measures associated with this condition to evaluate the application of EHR data to performance measurement.

The impact of diabetes on morbidity and mortality in the United States as well as its impact on health care costs make it a particularly important disease to manage. In 2010, 25.8 million people in the United States had diabetes and an additional 79 million Americans had prediabetes. Diabetes is the leading cause of kidney failure, nontraumatic lower-limb amputations, and new cases of blindness among adults in the
United States. It is a major cause of heart disease and stroke and the seventh leading cause of death in the United States.\textsuperscript{57} The American Diabetes Association estimates that the cost of diabetes and its complications was approximately $174 billion dollars in the United States in 2011.\textsuperscript{58}

There is substantial room for improving quality of care in patients with diabetes. Two in five persons with diabetes have poor LDL cholesterol control, one in three has poor blood pressure control, and one in five persons still has poor glycemic control.\textsuperscript{59} Saydah et al. report that less than 10% of patients with diabetes meet all three of these goals.\textsuperscript{60} Screening, diagnostic, and therapeutic actions can improve diabetes outcomes.\textsuperscript{61} Studies show that control of glucose levels, blood pressure, and cholesterol levels can dramatically delay or prevent the microvascular and macrovascular complications of diabetes.\textsuperscript{61} Despite this knowledge, the majority of patients with diabetes are not meeting their therapeutic goals.

Health care professionals can play a role in improving diabetes outcomes and adherence to standards of care is measurable. The American Diabetes Association Standards of Care in Diabetes position paper provides health care professionals with strategies to prevent acute complications and reduce the risk of long-term complications in diabetic patients.\textsuperscript{61} Many of these actions, such as ordering of periodic HbA1c and low-density lipoprotein (LDL) labs are measurable. In addition, it is possible to measure intermediate outcomes, such as blood pressure, HbA1c, and LDL values in diabetic patients.

The combination of the impact of diabetes on the United States, the ability to improve the condition and its outcomes, and the ability to measure these improvements, make it an ideal candidate for quality measurement.\textsuperscript{56} As a result, it is a disease frequently included in performance measurement initiatives. It is also one of the disease
areas being transitioned from claims-based performance reporting to EHR-based reporting. Diabetes was one of only three diseases for which PRQI originally permitted the use of EHR data for quality reporting when making the transition to EHR-based reporting in 2010. Table 1.2 presents the diabetes-related PQRI measures.

With the transition to EHR-based reporting, there have been a few studies exploring how to use EHR data for measuring quality of care in diabetes. Specifically, few studies have focused on what variables in the EHR should be used to identify the target population of patients with diabetes. Researchers have used different combinations of the following data elements to identify patients with diabetes: problem list diagnosis, abnormal plasma glucose levels, elevated hemoglobin A1c (HbA1c) levels, diabetes-specific medication prescriptions, encounter diagnoses, and diabetes-

<table>
<thead>
<tr>
<th>Measure</th>
<th>Description</th>
<th>Denominator (Target Population)</th>
<th>Numerator</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemoglobin A1c (HbA1c) Poor Control</td>
<td>Percentage of patients with diabetes mellitus who had most recent HbA1c greater than 9%</td>
<td>Patients with a diagnosis of diabetes: All patients with a documented diagnosis of diabetes at any time in the patient’s medical record. To be eligible for performance calculations, patients must have at least two face-to-face office visits with the physician, PA, or NP during the measurement period.</td>
<td>Patients with most recent HbA1c &gt; 9% during the measurement period.</td>
<td>Intensive therapy of glycosylated hemoglobin (A1c) reduces the risk of microvascular complications.</td>
</tr>
<tr>
<td>Low Density Lipoprotein (LDL) Control</td>
<td>Percentage of patients with diabetes mellitus who had most recent LDL level in control (less than 100 mg/dL)</td>
<td></td>
<td>Patients with most recent LDL &lt; 100mg/dL</td>
<td>Persons with diabetes are at increased risk of coronary heart disease (CHD). Lowering serum cholesterol levels can reduce the risk for CHD events.</td>
</tr>
<tr>
<td>High Blood Pressure Control</td>
<td>Percentage of patients with diabetes mellitus who had most recent blood pressure in control (less than 140/80 mm Hg)</td>
<td></td>
<td>Patients whose most recent blood pressure &lt;140/80mmHg</td>
<td>Intensive control of blood pressure in patients with diabetes reduces diabetes complications, diabetes-related deaths, strokes, heart failure, etc.</td>
</tr>
</tbody>
</table>
related phrases in free-text physician notes.\textsuperscript{9,62-64} As seen in table 1.3, the specific criteria for each of these data elements varies across studies. Depending on the study, different medications are categorized as diabetes-specific medications, different HbA1c and plasma glucose cut-offs are used, and the time periods for prescriptions, diagnoses, and labs vary.\textsuperscript{9,62,63} For example, while Tu et al. required only one elevated HbA1c, the Harris et al. and Tang et al’s teams required at least two.\textsuperscript{9,62,63} Further, the cut-offs for the HbA1cs ranges included a measure of 6.5, 7.0 or a measure “consistent with diabetes.”\textsuperscript{9,62,63} To date there is no standard algorithm for identifying patients with diabetes using the EHR.

<table>
<thead>
<tr>
<th>Citation</th>
<th>Gold Standard</th>
<th>Definition</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tu, et al, 2011\textsuperscript{62}</td>
<td>Manual record review</td>
<td>HbA1c &gt;7.0</td>
<td>49.4</td>
<td>99.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rx for one insulin or one oral diabetes medication</td>
<td>51.9</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HbA1c&gt;7.0 or Rx</td>
<td>64.9</td>
<td>99.9</td>
</tr>
<tr>
<td>Harris, et al, 2010\textsuperscript{63}</td>
<td>Patients with diagnosis of diabetes in the Ontario Diabetes Database (ODD)</td>
<td>≥2 elevated glucose tests OR insulin Rx OR ≥2 antidiabetes Rx OR problem list dx</td>
<td>75</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥2 elevated glucose tests OR insulin Rx OR ≥2 antidiabetes Rx OR HbA1c &gt;=6.5 OR problem list dx</td>
<td>73</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥1 elevated glucose tests OR insulin Rx OR ≥1 antidiabetes Rx OR HbA1c &gt;=6.5 OR problem list dx</td>
<td>72</td>
<td>99</td>
</tr>
<tr>
<td>Tang, et al, 2007\textsuperscript{9}</td>
<td>Manual record review</td>
<td>Problem list diagnosis or antidiabetic rx or two lab tests consistent with diabetes</td>
<td>97.6</td>
<td>99.6</td>
</tr>
</tbody>
</table>
1.6 Description of project

For the past decade, policy makers have emphasized the critical role that performance measurement has in improving the quality of health care in the United States. Hundreds of performance measures have been developed and tested using claims data.\(^1\) Recently, organizations such as CMS have begun to accept EHR data for their performance measures.\(^9,33\) However, research on the validity of using EHR data in performance metrics has not kept pace with the rapid transition to EHR-based performance measures. This document describes the findings of a project designed to investigate the validity of using EHR data for diabetes performance measures.

As described in section 1.4, one of the key elements to generating a valid measure of performance is being able to accurately identify the target population for the measure.\(^27\) We used two strategies to determine the best methods of identifying the patients with diabetes, the target population for diabetes performance measures. First, we conducted seventeen, one-on-one in-depth interviews with providers who use an EHR. The purpose of these interviews was to collect information on how providers enter diabetes diagnoses in the EHR and determine what motivates these documentation practices. During these interviews, providers described how they recorded diabetes diagnoses in the EHR and identified organizational factors that influenced their data entry behavior. Based on their data entry practices, providers gave recommendations as to how to use EHR data to identify patients with diabetes. Further, they identified how certain data entry and data processing factors could impact performance measures. These findings are presented in detail in Chapter 2.

To complement the data collected in these interviews, we used quantitative measures to evaluate the validity of different EHR-based methods for identifying patients with diabetes. This part of the project focused on the diabetes performance measures
listed in Table 1.2 that CMS recently transitioned to EHR-based reporting. We used EHR data from the Geisinger Health System (GHS), an integrated health care system in Pennsylvania. In 1996, GC began implementing its EHR in outpatient clinics. Installation of the EHR was complete in all ambulatory care clinics by 2001. We used EHR data on primary care patients to evaluate eight different methods of identifying patients with diabetes. We evaluated the specificity and sensitivity of these methods. We then evaluated each method to determine whether they selectively identify patients who are more or less likely to meet quality standards. The findings from this portion of the study are reported in Chapter 3. While previous studies have evaluated the sensitivity and specificity of different EHR-based methods of identifying diabetes patients, this project took the additional step of evaluating the potential methods for selection biases associated with EHR-based methods.9,62,63

In Chapter 4, we return to the data from our clinician interviews. During these interviews, we discovered that clinicians have some concerns about the unintended consequences of using EHR data for quality initiatives. In some cases, providers discussed hypothetical scenarios while in other cases providers gave real examples of consequences they had personally witnessed. Recognizing that this concern was not unique to one or two respondents, we pursued a line of questioning around this topic. Chapter 4 presents our findings about the unintended consequences of using EHR data for performance measures reported by providers. While there is extensive evidence documenting unintended consequences of using EHR data for clinical decision support and other applications, this is the first paper to report on the consequences of using EHR data in performance measurement.
Chapter 2

Organizational Influences on Data Entry Practices in the Electronic Health Record:
Implications for Performance Measurement

2.1 Background

As described in Chapter 1, one of the important considerations in developing a performance measure is how to identify the target population.\textsuperscript{27} Transitioning from claims to EHR-based performance measures requires a reevaluation of what data fields to use to identify the target population, as new and different data are available in the EHR. When selecting what data fields to use, it is critical to understand the origin of the data in the EHR. Specifically, it is important to know who enters the data and what factors influence their data entry practices, as these factors could influence quality scores.

The EHR offers a broader range of clinical fields than claims databases. In addition to diagnoses and procedure orders found in claims databases, the EHR includes detailed information on diagnostic test results, vital signs, social and family history, prescriptions, and physical examination findings.\textsuperscript{36} However, provider documentation habits can vary both in terms of how providers make notations in the record (e.g. ‘diabetes mellitus’, ‘DM’) and where in the record they enter information (e.g. past medical history, problem list).\textsuperscript{20} This variation could be due to a range of factors including workflow, billing processes, reimbursement policies, and user-interface with the EHR system.\textsuperscript{37-41}
Documentation practices can impact performance measures, in some cases leading to over- or underestimates of performance quality. In a recent study Parsons and colleagues reported that mammography rates in one medical practice were underestimated because the EHR-based performance algorithm would only count a mammogram if the test was ordered as a ‘procedure,’ and not a ‘diagnostic image.’ Therefore, physicians who labeled mammogram orders as a ‘diagnostic image’ received poorer performance scores, despite their adherence to the quality standard.\textsuperscript{20}

Parson and colleagues’ study is one of the only reports in the literature describing how physician data entry practices in the EHR directly influence performance measurement.\textsuperscript{20} To date, no studies have focused specifically on how data entry practices or other factors could impact the identification of the target population for performance measures. Despite this lack of critical knowledge, organizations are increasingly using EHR data to measure quality. CMS and other organizations have developed specifications to guide institutions as to how to use EHR data fields to target populations for their PQRI measures.\textsuperscript{33} However, it remains unclear what factors influence how data are entered into these data fields and whether or not these fields could be a source of bias in quality reporting.

2.2 Purpose

The purpose of this study is to get a better understanding of the EHR data fields used to identify the diabetes population for the denominator in diabetes performance measures and to identify the potential sources of bias associated with using these data fields. Currently 3 of the EHR-based PQRI measures administered by CMS are for diabetes care.\textsuperscript{34} We specifically focused on what factors influence how diabetes data are entered into the EHR and how these factors could impact performance measures.

2.3 Methods
2.3.1 Research Design

We conducted an exploratory qualitative study, using a grounded theory approach, to learn how clinicians document diabetes diagnoses in the EHR and what factors influence their documentation patterns. We conducted key informant interviews with clinicians working in the Geisinger Health System (GHS). As described in Chapter 1, GHS has a fully operational EHR system. As this area of research is relatively unexplored, we approached the interview data without prior assumptions and were attentive to emergent themes.

2.3.2 Setting and Recruitment

The study was conducted with primary care clinicians (physicians, physician assistants, nurse practitioners) employed by GHS. GHS is an integrated health services organization that includes 3 hospitals, 12 specialty care clinics, and 41 community practices clinic sites (primary care clinics with/without multi-specialty care and outpatient surgery). The practice sites are located across 31 of the 67 counties in Pennsylvania. GHS began implementing its EHR system in the outpatient setting in 1996, with installation complete in all ambulatory care clinics by 2001. Overall, GHS serves a patient population that is poorer, older, and sicker than the national average. Most patients have multiple chronic diseases, such as diabetes, high blood pressure, and lung disease.

Recruitment for our study occurred in 2011 and clinicians were eligible to participate if they were primary care providers and had been employed by the health system for at least six months. We used a maximum variation strategy to identify clinicians with a broad range of perspectives on EHR-based quality measures and diabetes care. We sought referrals from the quality improvement department of the health system to identify clinicians with a range of levels of participation in the system’s
quality improvement initiatives and EHR system and with a range of levels of performance based on the metrics obtained from the health system’s existing diabetes quality performance measures. We recruited participants via e-mail and reimbursed participants at a rate equivalent to the cost of an existing level 4 visit, determined by the health system, which was the same rate for all participants.  

2.3.3 Study Participants

Seventeen of thirty-seven clinicians (46%) contacted agreed to participate in the study. The remaining 20 clinicians did not actively decline participation, but failed to respond to the recruitment e-mail. After 15 interviews, investigators felt they reached saturation. The 17 clinicians were from 13 different clinic community practice sites within Geisinger. The clinic sites ranged in size and geographic location, with some clinics more than 130 miles apart from each other. Fourteen participants were primary care physicians (9 family medicine, 5 internal medicine), two participants were physician assistants and one participant was a nurse practitioner.

Participants reported seeing 36 to 450 patients per month. The number of patients seen was largely dependent upon whether respondents had other roles in the health system (e.g., administrative) in addition to patient care roles. On average, clinicians who participated in this study reported that nearly 20 percent of their patients had a diagnosis of diabetes. Six respondents indicated that, in addition to their clinical work, they had some role in quality improvement initiatives or EHR optimization projects for the institution.

2.3.4 Measures

We used a semi-structured interview instrument to guide our interviews that explored how clinicians enter data in the EHR when diagnosing and treating patients with a diagnosis of diabetes. (See Appendix A.) The interview guide included four
primary sections 1) Clinician’s professional history/background; 2) How does the diagnosis of diabetes get entered into the EHR? (e.g., What EHR fields are used?, Who enters the diagnosis?, When is the diagnosis entered?); 3) What factors influence documentation behavior?; and 4) What EHR data fields would you recommend for use in identifying patients with diabetes? The interview guide included a combination of open-ended questions and follow-up question probes to allow for flexibility during the interview. Clinicians were prompted to specifically discuss four types of EHR data that have been reportedly used to identify patients with diabetes using an EHR: Problem list diagnoses, encounter diagnoses, diabetes-related medication orders, diabetes-related lab results (i.e., hemoglobin A1c and glucose levels). The interviewer specifically asked the clinicians to explain how each of these types of EHR data were used and the advantages/disadvantages of using these data to identify patients with diabetes. All interviews were conducted one-on-one and administered over the telephone. Each interview lasted approximately 30 minutes.

2.3.5 Analysis

A team of two investigators coded the individual interview transcripts to identify broad themes across the data. The investigators used both inductive and deductive methods to code the data. Using the deductive approach, the investigators developed a coding frame based on the questions asked in the interview guide as well as from concepts supported in the literature. As investigators reviewed the transcripts, they also applied an inductive approach, identifying themes that emerged from the interviewees’ comments. Investigators defined codes for these emergent themes and the new codes were added to the original coding frame. All transcripts were then reread and recoded according to the modified coding frame. To ensure consistency, investigators met to
discuss the emergent themes and make final coding decisions. Final decisions about coding and themes were made when investigators reached consensus.

Both prior to analysis and throughout the coding process, investigators consulted available literature on diabetes quality measures and EHR data. Specifically, investigators reviewed papers that addressed what data fields in the EHR are typically used to identify diabetes patients for research and quality purposes. In addition, investigators accessed relevant data from the system's EHR to help validate their findings. For example, we reviewed the frequency in which patients with diabetes had the diagnosis documented in different parts of the EHR (e.g., the problem list, encounter diagnosis, etc.) We used the qualitative data analysis software program Atlas.ti (version 6.0) to support our coding and analysis process.

2.4 Results

In this chapter, we report on three themes that emerged during our discussions with respondents. First, clinicians identified two data fields, the problem list diagnosis field and the encounter diagnosis field, as the locations in the EHR where clinicians most frequently document a diagnosis of diabetes. Second, clinicians endorsed the use of a problem list diagnosis for identifying patients with diabetes for quality measures, while clinicians expressed concern that depending on an encounter diagnosis could result in over- or under-identifying patients. Third, while clinicians were not specifically asked about organizational influences on their documentation behavior, it emerged that organizational factors have an influence over how diagnosis data is entered into the EHR. The sections below provide additional details on these findings.

2.4.1 Documenting Diabetes Diagnoses

All respondents were asked where in the EHR they document a diagnosis of diabetes. For sixteen of the seventeen respondents (94%), the problem list was the first
data field mentioned. While encounter diagnosis was only mentioned first by one respondent, all respondents reported that they also record diabetes as an encounter diagnosis during encounters when treating a patient with diabetes. About one-third of respondents also mentioned adding the diabetes diagnosis to the Past Medical History. (Respondents were not specifically asked about Past Medical History fields).

2.4.2 Clinician Recommendations for EHR Fields to Use in Identifying Patients with Diabetes

Respondents were asked to comment on whether or not specific parts of the EHR would be useful in identifying patients with diabetes. As presented in Table 2.1, all clinicians agreed that using problem list diagnoses would be a good way to identify patients with diabetes. Most clinicians also agreed that finding patients with an encounter diagnosis of diabetes would be another good way to find all the patients with diabetes in the health system, but expressed some concerns regarding the potential for this method to lead to over and under-identification of patients with diabetes.

Clinicians unanimously endorsed the use of the problem list to identify patients with diabetes. Their endorsement was consistent with reports of the heavy use of the problem list among clinicians in the system. In addition, clinicians reported that they believed that diabetes was consistently and correctly entered on the problem list. When probed to think of an example when diabetes had been on the problem list incorrectly or had been incorrectly missing from the problem list, few clinicians could provide an example. One physician commented that he thought that diabetes had been missing from the problem list about three times over the course of his 16 years at the health system.

Clinicians’ support of using the problem list was specific for diabetes. Clinicians reported that while the problem list was typically used consistently for diabetes, it was
used less consistently for other diseases and for acute conditions. One respondent noted, “Your acute disease wouldn’t always be transferred to the problem list. You know, if you come in with a cellulites, you’re not going to transfer that over to the problem list. But if you have emphysema, diabetes, yes, they’re going to pull over to the problem list because that’s going to be an ongoing problem.”

The majority of respondents felt that using an encounter diagnosis to identify patients with diabetes would also be a good approach. However, unlike the clinicians’ endorsement of the problem list diagnosis as a way to identify patients with diabetes, the support for using the encounter diagnosis was mixed. Several clinicians noted the potential to over-identify and under-identify patients when using the encounter diagnosis.

Clinicians noted that the risk of under-identifying patients could occur because the encounter diagnosis was dependent upon a patient coming into the office for treatment of their diabetes. A clinician explained, “Let’s say that this particular patient doesn’t belong to me, they belong to another physician, and they happen to be seen for episodic visits, which are unrelated to a diagnosis of diabetes. Then diabetes could be theoretically be not dealt with that entire time that they’re absent from their primary care physician.”

Over-identification was attributed to clinicians using a diagnosis of diabetes in place of other diagnosis codes (e.g. impaired fasting glucose) when attempting to rule out a diagnosis of diabetes. One clinician noted, “There are some physicians who, they don’t think to put elevated sugar or something like that. So they are not officially diabetic, and you just use diabetic to quickly get the orders in.”
Table 2.1 Clinician Comments Regarding Documentation of Diabetes Diagnosis in the Problem List and Encounter Diagnoses Fields

<table>
<thead>
<tr>
<th>EHR Data</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem List</td>
<td>If it’s not on the problem list then usually that patient's not being followed for diabetes.</td>
</tr>
<tr>
<td></td>
<td>I think we’re fairly consistent here with how we use the problem list for diabetes.</td>
</tr>
<tr>
<td></td>
<td>If I was seeing a patient that I didn’t know and I wanted to get an idea of what their medical problems were, that’s (the problem list) where I’d look.</td>
</tr>
<tr>
<td></td>
<td>I know the EPIC team, when we’re doing quality assurance … we will generate a list from the problem list and its accurate.</td>
</tr>
<tr>
<td></td>
<td>There’s quite a variety. There are some that put every little thing on there. And there are others that only put real significant things on there. And there are others that the same thing in on there forever. You know, even though the acute problem is gone.</td>
</tr>
<tr>
<td></td>
<td>Yeah, I mean, I would say the problem list is always accurate and, if I was going to have a problem with the problem list it is that there are too many things on there. You know, people just put all sorts of extraneous stuff on there.</td>
</tr>
<tr>
<td>Encounter Diagnosis</td>
<td>Sometimes physicians use diabetes as a rule-out diagnosis in their claims based information. You know, they're ordering a glucose or they're ordering a glucometer for a patients and they'll put diabetes down. The patient really doesn't have diabetes, but to rule-out diabetes. It's not the way you're supposed to do it. You're supposed to enter hyperglycemia or something like that, but that's not in practice the way it's sometimes done. So we are very skeptical about a patient who only has one encounter level diagnosis of diabetes.</td>
</tr>
<tr>
<td></td>
<td>If a person's blood sugars are ranging high, instead of putting something like abnormal chemistry or impaired fasting glucose, people will just stick diabetes on there. Its almost kind of like a rule-out diagnosis but there is no rule-out diagnosis in EPIC</td>
</tr>
<tr>
<td></td>
<td>If they have impaired fasting glucose and they want diabetes-testing supplies to be paid for, you have to label them as having diabetes.</td>
</tr>
<tr>
<td></td>
<td>It depends upon your patient panel. If it’s an urgent care population, where people come and go, then you would under-diagnose. If it is a fairly stable panel, and people come back or they get the system involved multiple times, you'll eventually catch them.</td>
</tr>
<tr>
<td></td>
<td>You might miss a fair amount with that because they may be in for blood pressure, cholesterol, and not really there for the diabetes management. I tell you, we would probably put that on the encounter list that day, probably 70 to 80 to 90% of the time, but there would be times when they are in for other reasons. So, the problem list would probably be more accurate than the encounter list.</td>
</tr>
</tbody>
</table>
2.4.3 Organizational Factors Influencing EHR Data Entry

During the interviews, we asked clinicians to explain how and why they enter data into the problem list and encounter diagnosis fields. When explaining how and why they use these fields, it emerged that there are a number of organizational factors that drive how data are entered into these EHR fields. As presented in Table 2.2, internal quality performance programs, workflow, and leadership pressure all play a role in how these fields are used when documenting a diagnosis of diabetes.

2.4.3.1 Workflow

Clinicians described the set of processes, or workflow, that results in entering data in the EHR during a patient encounter. When discussing how encounter diagnoses of diabetes are entered into the EHR, respondents reported that nurses, physician assistants, nurse practitioners, and physicians enter the diagnoses codes. While nurses can enter a diagnosis code, they cannot authorize the final sign-off on the code.

In addition to manually entering the encounter diagnosis, it emerged during the interviews that the institution’s EHR system offers a number of options for having a diagnosis code automatically listed as an encounter diagnosis. According to the clinicians, all medication, lab, and procedure orders require a link to a diagnosis. When an order is linked to a diagnosis code, that code is automatically listed as an encounter diagnosis. For example, if a physician orders insulin during a visit and links the insulin to diabetes, diabetes will be listed as an encounter diagnosis for that visit.

Clinicians also described EHR tools, such as smartsets, that result in automatically creating an encounter diagnosis. A smartset is an EHR tool used to complete routine encounters all on one form (orders, diagnosis, etc.) Some respondents reported using diabetes smartsets to place orders, such as a hemoglobin A1c lab order. According to respondents, the diabetes smartsets automatically link the
<table>
<thead>
<tr>
<th>Factor</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Workflow</strong></td>
<td>If a nurse is ordering a medication or procedure for a patient, they will link that medication or procedure to an encounter diagnosis.</td>
</tr>
<tr>
<td></td>
<td>You have to correlate the test with the diagnosis so that you move that from your problem list into your encounter list and its addressed then and your orders are entered to that diagnosis.</td>
</tr>
<tr>
<td></td>
<td>As long as diabetes was associated with that medication, it would automatically come over.</td>
</tr>
<tr>
<td></td>
<td>Yeah, you can use smartsets and pre set up notes that make it very easy to document diabetic care.</td>
</tr>
<tr>
<td></td>
<td>Mostly using smartsets for foot exams, for ordering it and a lot of the orders come through smartsets through nursing. HbA1c is on it. So all of that’s pended before I even get in the room. You know, it’s like pended, ready for me to sign.</td>
</tr>
<tr>
<td></td>
<td>Most of my ordering is done from smartsets and I already have the diagnosis linked through the smartsets.</td>
</tr>
<tr>
<td></td>
<td>Well, any time your order something and it's associated with any one particular diagnosis it's on the diagnosis for that particular encounter. So, if I saw you today for depression, for example, and you happen to have diagnosis for diabetes for a medication that has been refilled, even if we didn't talk about it, diabetes is not the reason you're there, we are talking about depression, diabetes would be on the encounter visit that day.</td>
</tr>
<tr>
<td><strong>Quality Imprvmt.</strong></td>
<td>In the EHR our quality measures are driven off of the problem list. So, I tend to use the problem list to put in the diabetes diagnosis and then all the reminders start firing</td>
</tr>
<tr>
<td></td>
<td>With the bundles (quality program), usually there's a pressure to put the diagnosis in and you think about it more.</td>
</tr>
<tr>
<td></td>
<td>I think if there's a bundle (quality program) it really emanates from the problem list. I don't think we get a bundle unless it, you know, is on the problem list.</td>
</tr>
<tr>
<td></td>
<td>In general, I would say, since its part of the bundles (quality program) its there (on the problem list) more frequently.</td>
</tr>
<tr>
<td></td>
<td>To some degree, (physicians are using the problem list) more so than before the bundle. Definitely, more so than before the bundle.</td>
</tr>
<tr>
<td><strong>Organiz. Pressure</strong></td>
<td>Everybody within the system is very aware that the numbers are looked at. With all the Medicare issues and quality assurance out there, I think everybody is pretty much aware of it. They break it down according to clinics, they break it down according to providers, and share it with us.</td>
</tr>
<tr>
<td></td>
<td>I's one of the things that the system has made a priority so you're more focused on things that are out there in front where you are getting measured and looked at for.</td>
</tr>
<tr>
<td></td>
<td>We were very compulsive about putting our diagnosis in….we were told to do that.</td>
</tr>
<tr>
<td></td>
<td>There's a system level push to actually add the correct diagnosis to the problem list.</td>
</tr>
</tbody>
</table>
hemoglobin A1c order to a diagnosis of diabetes. Respondents also described best practice alerts, such as reminders for diabetes foot exams. These best practice alerts are sometimes linked to smartsets. Once again, some of these smartsets automatically link the resulting orders to diabetes, ultimately resulting in diabetes as an encounter diagnosis. Use of these EHR tools varied by clinicians. Some clinicians noted they frequently use smartsets while others reported never using smartsets. As a result of the automated entries of diagnosis codes, patients with diabetes whose clinician ordered a lab, medication, or procedure are more likely to have an encounter diagnosis of diabetes in the EHR system than a patient whose physician has not placed any of these orders. Alternatively, patients whose clinicians did not place an order linked to diabetes would be less likely to have an encounter diagnosis of diabetes.

2.4.3.2 Internal Quality Improvement Program

When discussing the problem list, it emerged that an internal quality improvement program for diabetes, referred to as the diabetes bundle, was driven off of the problem list. If a patient has a diagnosis of diabetes listed on their problem list, the patient will be included in the denominator for the diabetes bundle quality measures.

Because of this link between the quality improvement program and the problem list, clinicians indicated they were more conscientious about adding diabetes to the problem list saying, “because of the bundle, we try to make it look good.” An interview with the physician who developed the diabetes bundle indicated that the link between the bundle and the problem list was intentional. “We wanted to put them (physicians) in control of the accuracy of the data. And so they are the ones that populate the problem list and so they are the ones who tell me who their patients are with diabetes. And so it gives them the responsibility of making sure that the registry information is up-to-date and accurate.”
2.4.3.3 Organizational Pressure

Clinicians reported feeling pressure from leaders in their healthcare institution to keep their problem lists accurate and up-to-date. When asked how leadership exerted this pressure to maintain the problem list, most respondents cited the organization’s regular distribution of quality reports. In addition, respondents working in newer clinics noted that when the clinicians were integrated into the system, leadership emphasized the general importance of maintaining accurate and up-to-date problem lists. One respondent noted that the pressure to maintain the problem list has resulted in a problem list that is stronger than lists at other institutions. He explained, “So we have started to work with other organizations around the country, who are working with diabetes and have registries as well, and the biggest difference or the biggest debate is whether or not you use problem list data or not. Many EHR groups are not sure how accurate it is… and concerns come up over using problem list data.”

2.5 Discussion

The interviews with clinicians revealed that clinicians in this health system primarily use the problem list fields and encounter diagnosis fields for recording a diagnosis of diabetes. Clinicians supported the use of problem list diagnoses for identifying patients with diabetes for quality measures. While most of the clinicians also endorsed the use of encounter diagnosis, some clinicians felt that the use of this data field could result in over- and under-identification of patients with diabetes. In addition, during the interviews it emerged that there are organizational factors that influence both how the problem list and encounter diagnoses are entered into the EHR. In some cases, these factors could bias quality reports.

The clinicians’ descriptions of organizational factors that influence the EHR data
are consistent with Terris, et al’s conceptual framework for identifying potential sources of bias in secondary database analyses. Terris and colleagues identify health system-based factors as one of the primary potential sources of bias when using EHR data. Previous research has also reported that factors such as workflow and EHR user-interfaces, likely to be organization-specific, have an influence on how clinicians use the EHR.  

We identified three factors at the health-system level that influence how data are entered into the EHR fields. Specifically, system-specific workflow, health system priorities, and system-level quality improvement initiatives all influence the consistency with which diabetes is entered on the problem list and when diabetes is listed as an encounter diagnosis.

Some of these organization-specific factors may explain why the clinicians’ unanimous support for using the problem list for identifying patients with diabetes contrasts with experiences in other health systems. Literature describing the general use of EHR-based problem lists report that it is usually poorly maintained. Inter-rater reliability of diagnoses on the problem list has been described as moderate at best. The clinicians in our study report a number of system-level factors that could account for why the problem list in this organization may be uniquely complete. The system-level pressure to keep problem lists up-to-date as well as the direct link between the problem list and the diabetes quality reporting for the system’s internal quality reporting program appear to drive clinicians to be particularly conscientious about accurately recording diabetes on the problem list. Consistent with the success in this health system, work in other institutions has demonstrated the role that institutional interventions can have in improving clinician’s use of the problem list.

The same system factors that make the problem list a uniquely good source for identifying patients for performance measures at Geisinger, could make the problem list
a poor data field to use when comparing quality measures across healthcare systems. When developing algorithms for identifying patients with diabetes, organizations such as CMS need to consider the potential variability of the accuracy of EHR data fields across institutions.\textsuperscript{35,63,72} The organization we studied, according to the clinicians we interviewed, has a uniquely accurate and complete problem list, in part, due to system-specific factors. Healthcare systems without such factors could have less reliable data in their problem list. If these systems used the problem list to identify their patients with diabetes, they might identify only a select subset of their patients, potentially biasing their quality scores. This variation across institutions makes the development of a single EHR-algorithm for use across all health systems particularly problematic, creating a significant challenge for CMS and other organizations that are attempting to create a standard for all providers.\textsuperscript{35,72,76}

System-level factors influence how encounter diagnoses are generated as well. As with other EHR-based systems, labs, procedures, and medication orders linked to a diagnosis of diabetes automatically create a diabetes encounter diagnosis. Perhaps more unique to the institution under study, clinicians reported that the system has sophisticated automated workflows in place that ensure that a diabetes encounter is generated automatically when a diabetes-specific order, such as when a hemoglobin A1c lab, is ordered. The link between orders and encounter diagnoses has the potential to bias quality measures that use encounter diagnoses to identify the disease population. For example, one quality metric frequently reported is the percent of diabetes patients who had a hemoglobin A1c ordered. If a diabetes encounter is generated every time a hemoglobin A1c is ordered, that means a person whose physician ordered a hemoglobin A1c is more likely to be included in the denominator (identified diabetes population) than a person whose physician did not order a hemoglobin A1c. As a result,
the percent of diabetes patients who had a hemoglobin A1c ordered could be artificially inflated.

Tang and colleagues provide some quantitative evidence of the potential bias in using encounter measures to identify patients. When comparing patients with fewer than two diabetes encounter diagnoses to patients with at least two, only 68% of those with fewer than two encounters had a glycohemoglobin measured in the last year compared to 97% of patients with at least two encounters. One possible explanation for this disparity is the fact that orders, such glycohemoglobin orders, are directly related to the number of diabetes encounters. The potential for the encounter diagnosis field to bias quality measures is particularly alarming given that the encounter diagnosis field is already used in many of the existing quality measures for diabetes and other diseases.

2.5.1 Limitations

We identified two limitations of this study. First, the clinicians who agreed to participate in our study may have been more interested and knowledgeable about the use of EHR data for quality reporting than those who did not agree to participate. As a result, these clinicians may be more conscious of how they enter EHR data and be influenced by slightly different factors than clinicians who are not particularly concerned with EHR-driven quality measurement. We took several steps to minimize the potential for selectively recruiting these types of clinicians. First, we requested that the administration provide us with a recruitment list of clinicians with a range of knowledge and experience in EHR data and quality reporting. Second, we provided a financial incentive to participants with the goal of attracting physicians regardless of their interest in the study topic. Given our response rate of 46%, relatively high for clinicians, we believe we effectively minimized this potential source of bias.

A second limitation was that we confined our study to clinicians from one health
system. The health system we studied has some unique features. First, it is an integrated health system that employs 800 physicians. Second, it is an early adopter of the EHR and other health information technology. We cannot generalize our findings to all health systems. We suspect that physicians practicing at other institutions would be influenced by their own set of organization-specific factors, some of which could overlap with the factors we identified. Future studies should investigate data entry practices across a range of health systems in order to capture a broader understanding of the organizational factors that should be considered when developing performance measures.

2.6 Conclusion

The use of EHR data for measuring quality is expanding. With this expansion, organizations such as CMS are making an effort to develop specifications as to what EHR data fields to use to identify patient populations for quality reporting. Our results reveal that there are a number of organization-specific factors that influence how data are entered into the EHR. These factors have the potential to bias quality reporting and make comparisons across institutions problematic. Given the rapid adoption of EHR data for performance measurement, it is imperative we expand our understanding of factors that influence how EHR data is entered and understand how these influences can impact performance measures. Future research should explore organization-specific influences present in other health systems. In addition, researchers should explore the possibility of developing EHR algorithms for identifying disease populations that are less sensitive to organizational influences. Until we increase our knowledge about EHR data, EHR-based quality measures should be interpreted with caution.
Chapter 3
Using Electronic Health Record Data to Measure Performance in Diabetes Care:
Evaluating the Validity of Methods of Identifying the Target Population

3.1 Background

As discussed in previous chapters, the validity of a performance measure depends, in large part, on how the target population for the measure is selected.\textsuperscript{27} Performance measures for diabetes care are among the most common performance measures used and were among the measures first to use electronic health record (EHR) data sources.\textsuperscript{33,55} However, there is no standard for identifying patients with diabetes in the EHR. A lack of standard is problematic, as the method of identifying patients with diabetes could directly impact measures of performance. The purpose of this study was to investigate the validity of different methods of using EHR data to identify patients with diabetes for diabetes quality measures.

In 2010, The Centers for Medicare and Medicaid Services (CMS) allowed healthcare providers to use EHR data for 10 performance measures from the Physicians Quality Reporting Initiative (PRQI).\textsuperscript{33} Of these 10 performance measures, three were for diabetes care.\textsuperscript{33} Since that time, more organizations are proposing to measure quality in diabetes care using EHR data.\textsuperscript{1} While these organizations provide specifications as to how to identify diabetes patients in the EHR, their specifications are frequently vague, and fail to specifically identify which EHR data elements to use.\textsuperscript{1}
The literature reveals a wide range of EHR-based algorithms used to identify patients with diabetes for research and quality purposes. As presented in Table 1.3, researchers have used different combinations of the following data elements to identify patients with diabetes: problem list diagnosis, abnormal plasma glucose levels, elevated hemoglobin A1c (HbA1c) levels, diabetes-specific medication prescriptions, encounter diagnoses, and diabetes-related phrases in free-text physician notes. Studies that examined the sensitivity and specificity of different combinations of EHR data elements have reported some differences. However, comparing sensitivity rates across studies is challenging, given that the gold standard used for a diabetes diagnosis has not been consistent. To date there is no standard algorithm for identifying patients with diabetes in the EHR.

A lack of a standard method is particularly problematic in measuring quality, as different combinations of EHR data elements could identify different subsets of diabetes patients. Depending on the data elements used, the diabetes definition could selectively identify patients who perform better or worse on performance measures. For example, if you identify diabetes patients using a problem list diagnosis, and patients with a problem list diagnosis are more likely to have severe diabetes, you might selectively identify patients who score poorly on performance measures. Alternatively, if you identify diabetes patients using encounter diagnoses, and patients with clinical encounters are more compliant than other patients, you might selectively identify patients who perform better on performance measures. In other words, the data elements used to identify patients could be associated with quality performance, resulting in selection bias.

The potential for selection bias in methods of identifying diabetes patients for performance measures is problematic for two reasons. First, it makes comparison across providers and institutions difficult unless the same methods are used to identify patients.
patients. However, hospitals across the United States are at various stages of EHR implementation and do not have the same access to electronic clinical data. As a result health systems early in the adoption process may be forced to use a more limited number of data elements when identifying a disease population in the EHR, making comparison of quality performance challenging. Second, the potential for selection bias opens the door to manipulation of performance measures. Studies of pay-for-performance programs have reported evidence of providers excluding subsets of patients from denominators in quality reporting in order to inflate performance scores. If certain data elements are associated with better performance, making small changes to the EHR algorithm used to identify patients could selectively identify healthier patients and artificially inflate performance scores.

3.2 Purpose

The purpose of this study was to examine the validity of different approaches to using EHR data to identify patients with diabetes for quality measures. We hypothesized that different combinations of EHR data elements would identify different subsets of diabetes patients. Furthermore, we predicted that the data elements might selectively identify patients who perform better or worse on quality performance.

3.3 Methods

3.3.1 Setting and Recruitment

We conducted a data pull from the EHR of Geisinger Health System (GHS). GHS serves approximately 450,000 primary care patients. The health system installed the EHR in the outpatient clinics in 1996 and installation was complete in all ambulatory care clinics by 2001. The EHR was introduced in the inpatient setting in 2008. We extracted EHR data on all primary care patients, 18 years of age and older, who had at least one outpatient encounter prior to 2007 and at least two outpatient encounters in 2009. These
criteria helped ensure all eligible patients had been in the system for at least two years, and that patients were still getting their primary care from this health system during the study period (January – December 2009). We extracted demographic, vital signs, diagnosis, medication, and laboratory data on individuals meeting these criteria.

### 3.3.2 Diabetes Definitions

We created eight different definitions for diabetes, described in Table 3.1, using different combinations of the following EHR data elements: problem list diagnoses, encounter diagnoses, diabetes-specific medications, and elevated HbA1c levels. Definitions 1 through 4 require the presence of one specific data element in the record (e.g. Definition 2: Presence of problem list diagnosis). The remaining four definitions are more sensitive, and classify patients with diabetes if they have any one of two, three, or four of these data elements (e.g. Definition 8: Presence of problem list diagnosis OR encounter diagnosis OR medication OR elevated HbA1c level.) These eight definitions were derived from the CMS specifications for identifying patients with diabetes as well as other definitions used in previous literature. In addition, we received input from providers in the health system that use the EHR to manage their patients with diabetes. The data were automatically extracted from the EHR and provided to investigators in a SAS database.

We used a manual medical record review for our gold standard definition of a diabetes diagnosis. A trained, non-physician reviewer analyzed a subset of medical records randomly selected from the eligible primary care population. The reviewer categorized a patient as diabetic if, during the manual review, he/she found a diagnosis of diabetes on the problem list, found a diagnosis of diabetes associated with a medication or lab order, or found evidence of diabetes in the free-text notes documented by physicians in the Progress Note section of the EHR. When reviewing the free text
notes, the reviewer looked for phrases indicative of a diabetes diagnosis, including “diabetic foot ulcer” and “diabetic retinopathy.”

<table>
<thead>
<tr>
<th>Def. #</th>
<th>Data Element</th>
<th>Data Element Description</th>
<th>Definition Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Encounter diagnosis</td>
<td>Encounter diagnoses consist of a list of all conditions co-existing at the time of the encounter that affect the treatment received or length of stay. A condition of sufficient significance to warrant inclusion for investigative medical studies.</td>
<td>At least one diabetes ICD-code associated with at least two encounters between January 1, 2007 and December 31, 2009: 250.xx, 362.01-362.07, 366.41, 648.01-648.04.</td>
</tr>
<tr>
<td>2</td>
<td>Problem list diagnosis</td>
<td>The problem list constitutes a master list of all a patient’s diagnoses. This list includes clinical problems, a diagnosis summary and stressor exposure, an ongoing list of clinically significant health status events, resolved and unresolved, in a patient’s life.</td>
<td>At least one diabetes ICD-9 code on the problem list: 250.xx, 362.01-362.07, 366.41, 648.01-648.04.</td>
</tr>
<tr>
<td>3</td>
<td>Lab value: Elevated HbA1c</td>
<td>Documentation of theHbA1c test results from the clinical laboratory.</td>
<td>At least two HbA1c levels ( \geq 6.5 ) (on separate days)</td>
</tr>
<tr>
<td>5</td>
<td>Meet criteria for definitions 1 OR 2</td>
<td>See descriptions for definitions 1 and 2.</td>
<td>See criteria for definitions 1 and 2.</td>
</tr>
<tr>
<td>6</td>
<td>Meet criteria for definitions 1 or 2 or 3</td>
<td>See descriptions for definitions 1, 2, and 3</td>
<td>See criteria for definitions 1, 2, and 3.</td>
</tr>
<tr>
<td>7</td>
<td>Meet criteria for definitions 1,2 or 4</td>
<td>See descriptions for definitions 1, 2, and 4</td>
<td>See criteria for definitions 1, 2, and 4.</td>
</tr>
<tr>
<td>8</td>
<td>Meet criteria for definitions 1, 2, 3, or 4</td>
<td>See descriptions for definitions 1-4.</td>
<td>See criteria for definitions 1-4.</td>
</tr>
</tbody>
</table>

* Patients only prescribed metformin, thiazollidinidine, or exenatide needed to also meet criteria for at least one of definitions 1-3 as these medications are indicated for diseases other than diabetes.

3.3.3 Performance Measures

To date, CMS has provided specifications for three diabetes quality measures: HbA1c control, blood pressure control, and cholesterol control. Per the CMS
specifications, patients with diabetes whose most recent HbA1c level was less than or equal to 9.0 were considered under control; patients whose most recent blood pressure was below 140/80 mmHg were considered under control; and patients whose most recent low density lipoprotein (LDL) was less than 100mg/dL were considered under control. Additional details are provided in Table 1.2. CMS measures quality over the calendar year. For our analysis, we looked at the last of these values reported in 2009.

We calculated the quality measures for the three CMS diabetes quality measures: blood pressure, LDL, and HbA1c. When calculating the quality scores, CMS specifies that the denominator include all patients meeting the diabetes criteria, whether or not they had a record of the clinical test (i.e., blood pressure reading, LDL lab result, HbA1c lab result) in their medical record for the year being studied. This strategy prevents organizations from artificially inflating their scores by failing to measure patients least likely to meet the CMS quality requirements. The numerator includes all the patients whose clinical outcome measure met the quality threshold defined by CMS. For example, to calculate the LDL quality score for definition 1, one would divide the number of patients meeting the criteria for definition 1 by the number of patients meeting the criteria for definition 1 whose last LDL measure in 2009 was below 100.

3.3.4 Analysis

We extracted data from the EHR on the primary care patients meeting our criteria (at least 18 years of age, at least 2 visits in 2009, at least one visit before January 1, 2007). Next, we determined which patients met the criteria for diabetes, according to each of the eight definitions. Based on the definition criteria, patients could be defined as a diabetic by anywhere from 0 to all 8 of the definitions. We estimated the prevalence of diabetes for each of the diabetes definitions and for the gold standard. To calculate prevalence, we used the number of patients identified by the definition as the numerator
and the total number of primary care patients meeting the general inclusion criteria as the denominator.

We calculated the sensitivity and specificity of each of the 8 definitions by comparing them to the gold standard manual medical record review. For this part of the analysis, we looked at a randomly selected subset of 499 of the eligible primary care patients. Previous studies that have used the EHR to identify patients with diabetes have resulted in false positive and false negative rates of less than 5%. To estimate the false positive rate in each denominator with a 95% confidence interval (CI) of +/-10%, we determined we needed to apply the gold standard medical record review to at least 34 patient records identified in each of the denominators, based on the conservative estimate of a 10% false positive rate. Similarly, to estimate the false negative rate with a 95% CI of +/-10%, we needed to identify 34 records of patients with diabetes using the gold standard record review. Based on the prevalence of diagnosed diabetes among individuals 20 years of age and older, we determined we would need to review the records of approximately 500 individuals to identify 40 patients with a diagnosis of diabetes.

We calculated the proportion of patients identified by each of the eight definitions who met the CMS quality standards. To determine whether a definition selectively identified patients with better or worse performance scores, we used the definition with the highest combination of sensitivity and specificity to identify the “true diabetics” in the primary care population. We then used the chi-square test to determine whether quality performance of the “true diabetics” identified by each of the other seven definitions differed from the quality performance of the “true diabetics” excluded by each of the definitions. In comparing quality scores, we controlled for age and diabetes severity using logistic regression analysis.
3.4 Results

We identified 125,102 primary care patients who met the initial inclusion criteria. About 17 percent of these patients met the criteria for at least one of the eight definitions for diabetes, as shown in Table 3.2. The prevalence estimate of diabetes in the primary care population was highest (17.2%) when the minimum criteria for diabetes was whether a patient had either a problem list diagnosis, encounter diagnosis, diabetes medication order, or elevated HbA1c levels (definition 8), and lowest (12.2%) when the minimum criteria was having at least two elevated HbA1cs (definition 3). Among those definitions that required the presence of one specific data element to identify patients with diabetes (definitions 1-4), defining diabetes as having at least two encounters with diabetes (definition 1) resulted in the highest prevalence rate (16.3%).

Sensitivity and specificity analyses were conducted on a subset of 499 patients randomly selected from primary care patients. The manual record review (gold standard) identified 80 of these individuals as having a diagnosis of diabetes. Definitions for diabetes that used encounter diagnoses (definitions 1, 5, 6, 7, 8) had sensitivity over 95% (Table 3.2). The definition that required an order for a diabetes-specific medication (definition 4) had a sensitivity of 83.8% and the definition that required a problem list diagnosis (definition 2) had a sensitivity of 75%. Requiring at least two elevated HbA1c levels (definition 3) resulted in the lowest sensitivity (65%). Of the 80 patients who had a diagnosis of diabetes according to the gold standard definition, requiring that the patient have two encounters associated with diabetes (definition 1) identified all but 3 patients as having diabetes. Of these three, one of the patients not identified was not diagnosed with diabetes until December of 2009, the last month in the study period.
All definitions had specificity over 99.4%. Two patients of the 419 who did not have diabetes, per the gold standard, had at least two encounter diagnoses of diabetes (definition 1). One of these patients also had a diabetes-specific medication ordered (definition 4). While these patients were documented as at risk of getting diabetes in the progress notes, a diagnosis of diabetes was ruled out, per the progress notes, as of the end of the study period.

Approximately 65% of patients identified by each of the definitions met the CMS quality standard for blood pressure control, as shown in Table 3.3. The proportions ranged from 64.8% when definition 3 was used, to 65.5% when definition 2 was used.
There was a slightly larger range in the proportion of patients meeting the LDL quality standard. The proportion of patients meeting the standard ranged from 67.1% identified by definition 8, to 71.3% identified by definition 3. For each definition, more than 80% of patients met the quality standard for HbA1c control. The proportions ranged from 82.9% among those identified by definition 8, to 85.1% among those identified by definition 3.

Table 3.3 Proportion of Patients Identified by Diabetes Definitions Meeting CMS Quality Standards: Patients Identified by Definitions and Patients Missed by Definitions

<table>
<thead>
<tr>
<th>Definition</th>
<th>Number of patients</th>
<th>Percent meeting BP Standard</th>
<th>Percent meeting LDL Standard</th>
<th>Percent meeting HbA1c Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 At least 2 encounter diagnoses</td>
<td>20431</td>
<td>65.1%</td>
<td>64.2%</td>
<td>68.7%</td>
</tr>
<tr>
<td>2 Problem List</td>
<td>16322</td>
<td>65.5%</td>
<td>63.7%</td>
<td>70.2%</td>
</tr>
<tr>
<td>3 Two HbA1c &gt;6.5 (on separate days)</td>
<td>15281</td>
<td>64.8%</td>
<td>65.8%</td>
<td>71.3%</td>
</tr>
<tr>
<td>4 Medication</td>
<td>18399</td>
<td>65.3%</td>
<td>63.9%</td>
<td>68.4%</td>
</tr>
<tr>
<td>5 2 Encounter dx OR Problem List</td>
<td>20570</td>
<td>65.1%</td>
<td>64.6%</td>
<td>68.4%</td>
</tr>
<tr>
<td>6 2 Encounter dx OR PL OR A1c</td>
<td>20881</td>
<td>65.0%</td>
<td>67.7%</td>
<td>68.1%</td>
</tr>
<tr>
<td>7 2 Encounter dx OR PL OR Med</td>
<td>21418</td>
<td>65.1%</td>
<td>59.9%</td>
<td>67.2%</td>
</tr>
<tr>
<td>8 At least one of the above</td>
<td>21555</td>
<td>65.1%</td>
<td>67.1%</td>
<td>82.9%</td>
</tr>
</tbody>
</table>

p<0.0001 – Controlled for age and diabetes severity
p<0.05 – Controlled for age and diabetes severity
p<0.01 – Controlled for age and diabetes severity
Based on the combination of high sensitivity and specificity, we used definition 8 to identify the “true diabetics” in the population. Definition 8 was also the most inclusive definition, classifying a patient as diabetic if they met the criteria for any of the first seven definitions. Each of definitions 1 through 7 missed a portion of the true diabetics. (See Table 3.2) In all but one case, the patient group that the definition classified as diabetic performed better on the LDL and HbA1c quality measures than the patient group the definition excluded. This was particularly true for definitions that classified patients as diabetic if they had 2 or more encounter diagnoses of diabetes (definitions 1, 5, 6, and 7).

Approximately 69% of diabetics identified by definition 1 (2+ encounter diagnoses) met the LDL standard, compared with 37.8% of diabetics who were not classified as diabetic by definition 1 (p<0.0001). Close to 85% of patients who met the criteria for definition 1 met the HbA1c standard compared with 52.2% of patients who did not meet the definition criteria. Over 84% of patients classified as diabetic by definition 5 (2+ encounter diagnoses OR problem list diagnosis) met the HbA1c quality standard, compared to 54.2% of patients excluded by the definition criteria (p<0.0001). All differences remained significant after controlling for age and diabetes complication severity.

3.5 Discussion

While organizations are migrating to the use of EHR data to measure quality in diabetes care, there is no standard for identifying patients with diabetes using EHR data. In our study we evaluated the validity of eight EHR-based algorithms to identify primary care patients with diabetes. We then studied whether the combination of EHR data elements used to identify patients was associated with performance on quality measures for diabetes. We determined the combination of EHR data elements used to identify a
diabetes population have varying degrees of sensitivity. In addition, in some cases, the data elements used to define a diabetes population selectively identify patients with better quality performance. As a result, the use of some EHR algorithms for identifying patients with diabetes could artificially inflate quality performance scores. In addition, comparing performance across health systems could be problematic in the absence of a standard method for identifying diabetes patients.

The sensitivity of the diabetes definitions we used in our study ranged from 65% to 98.8%, while all of the definitions had specificity above 99%. Our findings are similar to studies that have examined the sensitivity and specificity of similar EHR-based definitions for diabetes. Overall, the literature reports that most EHR-based definitions for diabetes, regardless of the combination of data elements used, have high specificity (over 97%).\textsuperscript{9,62,63} Of the eight definitions we evaluated, the definition with the most sensitive criteria (definition 8: either problem list diagnosis, 2+ encounter diagnosis, medication order, or 2+ elevated HbA1cs) had a sensitivity of 98.8%. Tang et al. used a similar definition and reported a sensitivity of 97.6%.\textsuperscript{9} Tu examined the sensitivity of requiring 1 HbA1c lab over 6 and reported a sensitivity of 76.6%.\textsuperscript{62} Consistent with Tu's findings, when we required two labs over 6.5, we found a relatively lower sensitivity (65%).

The quality performance in the population we studied was slightly better compared with other reports of blood pressure, LDL, and HbA1c control in patients with diabetes. However, we expected better performance in the population we studied, as all patients in the health system with a diabetes diagnosis are automatically included in a diabetes quality improvement program. Approximately 65% of patients with diabetes in the health system we studied, regardless of the definition used, met the quality standard for blood pressure. AHRQ reports a slightly lower number, 61%, among patients with
similar ethnic backgrounds as our population.\textsuperscript{86} LDL performance in our population ranged from 67\% to 71\%, slightly higher than a report from the National Committee for Quality Assurance (NCQA) of 60.4\% of diabetic patients.\textsuperscript{87} The HbA1c control was higher in our population compared to others. While more than 80\% of patients identified by each of our definitions met the HbA1c quality standards, NCQA reports that 68\% of patients with diabetes meet the HbA1c standard.\textsuperscript{87}

The definition used to identify patients with diabetes did not appear to have a large impact on the quality performance for blood pressure, LDL, or HbA1c measures. Patients identified across the eight definitions scored within 1.7 percentage points of each other on the blood pressure measure, within 4.1 percentage points of each other on the LDL measure, and within 2.2 percentage points of each other on the HbA1c measure. However, we determined that the EHR criteria used in each of the definitions did selectively identify patients who had better quality performance scores. Specifically, definitions that classified patients as having diabetes if they had at least two encounter diagnoses associated with diabetes (definitions 1, 5, 6, and 7) selected patients with better quality performance. The true diabetics who were excluded by these definitions (i.e., patients with fewer than two encounters) had significantly lower performance scores. As noted in Chapter 2, Tang reported a similar finding when comparing the use of one encounter diagnosis versus two encounter diagnoses to identify patients with diabetes for quality measures. Patients meeting the criteria for two encounters performed significantly better than patients who had fewer than two encounters (Blood pressure control- 61\% vs. 45\%, \( p<0.001 \); HbA1c testing- 97\% vs. 68\%, \( p<0.001 \)).

There are a number of possible explanations for the quality disparity between those with at least 2 encounter diagnoses and those without. First, as we learned during the clinician interviews, encounter diagnoses for diabetes are generated (either manually
or automatically) in the EHR system we studied whenever a provider orders a diabetes related lab measure, such as HbA1c. Therefore, a patient with an encounter diagnosis is more likely to have had an HbA1c test result in the record than a patient without an encounter diagnosis for diabetes. As a result, when the definition for diabetes requires the presence of two or more encounters with a diagnosis of diabetes, one is selectively identifying patients more likely to have had the HbA1c test. Per the CMS quality specifications, those without an HbA1c test in the record are considered to have not met the quality measure. By selectively identifying patients who are more likely to have had the test, you are selectively identifying a patient population that will likely perform better on the quality performance measures for HbA1c.

Alternatively, a lack of encounter diagnosis could be a signal that patients are having their diabetes managed primarily outside of the health system. As a result, measures for LDL, blood pressure, and HbA1c that are absent from the EHR under study, may have been conducted at another institution and never transferred over to the EHR. As a result of the missing tests, these patients were considered to have failed to meet the quality standard. We tried to ensure that the patients we studied used the health system for their care by excluding patients with fewer than two visits to a primary care provider in the system in 2009. In addition, when we excluded patients who did not have test results in the record from our analysis of quality performance (analysis not shown), there was still a significant difference in performance between those with two or more encounters and those with less than two.

The association between methods of identifying patients and quality performance suggests that there is potential for manipulation of quality performance reports based on the criteria used. Making changes to the EHR algorithm used to identify patients could select a healthier diabetic population than actually exists. For example, excluding
patients who do not have at least two encounter diagnoses may exclude patients with poorer quality performance scores. Studies of pay-for-performance programs have reported evidence of providers excluding subsets of patients from denominators in quality reporting in order to inflate performance scores.\textsuperscript{81-83, 88,89} Chen and colleagues specifically examined a diabetes pay-for-performance program and reported that a group of providers improved their performance by excluding older patients and patients with more severe conditions.\textsuperscript{81} With the transition from paper records to EHR systems, complex algorithms can be applied to permit careful selection of those patients who are more likely to perform well with respect to quality performance measures. To minimize the potential for health systems to manipulate inclusion criteria for performance measures, quality organizations have to be very specific in defining what data elements the organization is required to use to identify patients with diabetes and what exclusions are appropriate.

These findings have important implications when comparing quality performance across organizations. Most health systems in the United States do not have complete EHR systems.\textsuperscript{13,14} The Health Information and Management System Society (HIMSS) reports that in the third quarter of 2011, adoption rates in more than 5000 hospitals in the United States and Canada ranged from no part of an EHR system installed (9.6% of hospitals surveyed) to the installation of a complete EHR (1.1% of hospitals surveyed).\textsuperscript{14} On average, hospitals fell somewhere in between these two extremes, with hospitals having different parts of an EHR system in place.\textsuperscript{14} Therefore, the majority of hospitals is limited to using only those data elements that exist in a partial EHR system when identifying patients with diabetes. Our study shows that it is essential that health systems be explicit about how they identify their diabetes population when reporting quality performance. To fairly compare performance across organizations, the definitions used
and the access to certain parts of the EHR have to be considered, as these factors can result in excluding low- or high-performers.

While in recent years there have been a few studies investigating the sensitivity and specificity of EHR-based methods to identify the target population for diabetes measures, Tang, et al. conducted the only other study, to our knowledge, to assess these methods for selection bias. However, Tang et al. only assessed selection bias for one method of identifying patients, the use of two or more encounters. In addition, the Tang, et al. study was confined to a Medicare population. Our study evaluated and reported on selection bias for seven different EHR-based methods for identifying patients with diabetes. Specifically, we looked beyond encounter counts and evaluated methods that used problem list diagnosis, medications, and lab results. In addition, ours is the first study to evaluate selection bias in a general adult primary care population.

Understanding the association between EHR criteria and quality performance can help health systems target populations who are more likely to have poor LDL and HbA1c control. Our findings reveal what factors are associated with higher or lower odds of meeting quality standards. This is critical information, as less than 10% of patients with diabetes meet their HbA1c, blood pressure, and cholesterol goals. Control of glucose levels, blood pressure, and cholesterol levels can dramatically delay or prevent the microvascular and macrovascular complications of diabetes and quality programs have successfully improved control of these factors in diabetes patient populations. Tailoring programs to high-risk individuals can improve the efficiency of quality measures and their resultant quality improvement interventions.

3.5.1 Limitations

We identified two limitations to our study. First, we only studied EHR data from one health system. As EHR data is, in part, dependent on system policies and provider
data entry, it is possible that the way in which diabetes data is stored in the EHR can differ across institutions. In addition, the health system we studied has a complete EHR, while most health systems across the countries have not yet adopted or have only adopted partial EHR systems. Finaly, this health system has a diabetes quality program in place that automatically reminds providers to comply with diabetes quality standards for diabetes patients. Given the characteristics of this health system, our findings may not be generalizable across other health systems, such as those with incomplete EHRs or those without internal quality improvement programs. An important next step is to complete this same analysis in other health systems with different EHR systems and those at different stages of EHR adoption.

In addition, we only verified the diagnosis of diabetes for a subset of the patients we studied. We restricted our manual chart review to 499 patients. For our analysis of quality measures, we wanted to leverage our access to the large number of patients with diabetes in the health system’s medical record; however, we were unable to manually review the charts of more than 20,000 patients. Nonetheless, we are confident, based on our specificity data and data from previous studies, that we had a minimal number of patients in our analysis who were misclassified as having diabetes.

3.6 Conclusion

The availability of detailed clinical data in the EHR makes it a very appealing data source for measuring and monitoring quality performance. However, careful consideration has to be given as to how patients assessed with respect to these quality measures will be identified using EHR data. We found that the use of different data elements to identify patients with diabetes can result in selectively identifying patients more likely to meet quality standards. This is problematic as intentional or unintentional manipulation of EHR algorithms for identifying patients with diabetes could result in
excluding patients who are more or less likely to meet quality thresholds. As a result, quality organizations need to carefully craft EHR criteria for identifying patients with diabetes and provide detailed specifications as to how organizations are required to identify patients for quality reporting.
Chapter 4

The Unintended Consequences of Leveraging Electronic Health Record Data for Quality Improvement

4.1 Background

The American Medical Informatics Association (AMIA) has acknowledged the importance of understanding the unintended consequences of health information technology (HIT). In 2009, the organization dedicated its annual policy meeting to considering this topic and has since recommended that researchers conduct additional research into the causes, manifestations, and mitigation of unintended consequences of HIT implementation. The use of electronic health records (EHRs) for quality improvement programs is a rapidly growing application of HIT. However, little attention has been paid to the potential for unintended consequences of this use of EHR data.

The Institute of Medicine (IOM) lists the ability to report data, including quality data, as one of the key functional elements of an EHR system. Consistent with IOM’s definition of an EHR, the Centers for Medicare and Medicaid (CMS) requires that clinicians use EHR systems to report quality measures in order to meet the meaningful use criteria. Already CMS is permitting the use of EHR data for its pay for performance measures. In addition to quality reporting, healthcare organizations are using EHR data in other quality improvement initiatives that provide automated clinical decision support to physicians and automated updates to the medical record.
There is ample documentation of unintended consequences associated with such HIT applications as bar code medication technology, computerized physician order entry systems (CPOE), and electronic prescribing.\textsuperscript{47-52,54,96,97} Frequently the focus of the literature has been on the negative consequences, which range from inefficient workflows to the extreme of patient death.\textsuperscript{47,49-51} HIT is particularly vulnerable to these consequences in the early phases of implementation.\textsuperscript{98} Deriving performance measures from EHR data is a relatively new way to use this technology and we know very little about the unintended consequences.\textsuperscript{99-101} This is particularly concerning as CMS incentives for using EHR data in quality reporting will likely result in an increase in the number of providers who use EHR data to measure quality.\textsuperscript{5} Without understanding the potential consequences, these providers do not have the critical information needed to prevent and address such consequences.

The purpose of this study is to gain a better understanding of the potential unintended consequences of using EHR data to identify patients for quality improvement programs. To obtain this understanding, we conducted open-ended interviews with clinicians working in a health system that uses EHR data for performance measurement and other quality initiatives. We specifically focused on the consequences of using EHR algorithms to identify patients with diabetes for performance measures, as diabetes is among the disease areas that CMS currently permits the use of EHR data for quality reporting.\textsuperscript{33}

4.2 Methods

4.2.1 Research Design

As described in Chapter 2, we conducted a qualitative study, using a grounded theory approach to gain clinicians’ perspectives about the process of using EHR data to identify patients with diabetes for performance measurement.\textsuperscript{66} We conducted key
informant interviews with clinicians working in a health system with a fully functional electronic health record system. Throughout the data collection phase, investigators met to discuss themes that emerged from the interviews. When investigators came to consensus about emergent themes, new questions were added to the interview guide to capture data on these emergent themes. One of the themes that emerged during the interviews was clinician concern about the unintended consequences of using EHR data for quality reporting. The next section describes how we captured and analyzed the data around this theme.

4.2.2 Setting and Recruitment

As stated in Chapter 2, the study was conducted with primary care clinicians employed by Geisinger Health System (GHS). Clinicians were eligible if they were physicians, physician assistants, or nurse practitioners in primary care (family medicine and internal medicine) that had worked at the health system for at least six months prior to recruitment.

We used a maximum variation strategy to identify clinicians with a broad range of perspectives on EHR-based quality measures. To operationalize this strategy we contacted the institution's quality improvement department and requested a list of forty clinicians with a range of involvement with the system's existing EHR-based quality measures. This list included clinicians who held lead roles in quality improvement initiatives as well as clinicians with little to no involvement in quality improvement projects. We recruited participants via e-mail and reimbursed participants at the equivalent of the cost of an existing level 4 visit.

4.2.3 Study Participants

Seventeen of thirty-seven clinicians (46%) contacted agreed to participate in the study. Investigators felt they reached saturation on the topic of unintended
consequences after 15 interviews. Fourteen participants were primary care physicians (9 family medicine, 5 internal medicine), two participants were physician assistants and one participant was a nurse practitioner. Participating clinicians were drawn from thirteen different outpatient clinic sites across the health system. Participants reported treating 36 to 450 patients per month. On average, participants estimated that 20 percent of these patients had a diagnosis of diabetes. Approximately one-third of participants (6 of 17) reported they have a role in quality improvement or EHR improvement initiatives at the health system.

4.2.4 Measures

As described in Chapter 2, we used a semi-structured interview guide to capture information about how clinicians enter data in the EHR when diagnosing and treating patients with a diagnosis of diabetes. The interview consisted of four primary sections 1) Clinician’s professional history/background; 2) How does the diagnosis of diabetes get entered into the EHR? (What EHR fields are used, Who enters the diagnosis, When is the diagnosis entered); 3) What factors influence documentation behavior; 4) What EHR data fields would you recommend for use in identifying patients with diabetes? (Why, Why not?). The interview guide was a combination of open-ended questions and probes to allow for flexibility during the interview. All interviews were audio-recorded and transcribed.

As themes emerged during the data collection period, the interviewer would add questions to the interview guide for future interviews. One of the earliest themes that emerged was respondents’ concerns about the unintended consequences of using an EHR algorithm to identify patients with diabetes for quality reporting. As a result, after the first few interviews, the interviewer added a question to the interview guide regarding the potential for unintended consequences: “Do you have any concerns about the
The purpose of this paper is to present our findings on clinicians’ concerns about these unintended consequences.

4.2.5 Analysis

Prior to analysis, investigators consulted available literature on the use of EHR data in diabetes measures. Investigators developed an initial coding frame based on the literature and on the categories of questions in the original interview guide. Consistent with the grounded theory approach, throughout the coding process investigators met to discuss emergent themes and develop new codes for these emergent themes. As new codes were developed, investigators would recode interviews with the updated coding scheme. To ensure consistency, investigators met to discuss coding decisions. Final decisions on coding and emergent themes were made when investigators came to consensus.

We used the qualitative data analysis software program Atlas.ti (version 6.0) to code and analyze the interview transcripts. For this chapter we specifically focused on the theme that emerged around the unintended consequences.

4.3 Results

The section below highlights four types of unintended consequences that described in the interviews. Not all of the unintended consequences were negative. Further, while respondents did identify some negative consequences, respondents overwhelmingly supported the use of EHR data to facilitate quality reporting. The major themes of unintended consequences that emerged were: 1) Improved documentation in the EHR; 2) Negative impact on care process; 3) Insurance coverage issues; and 4) Unnecessary patient anxiety. Below we describe these themes in greater detail, with supporting quotes presented in table 4.1.
4.3.1 Improved Documentation

When discussing how to use EHR data to identify patients with diabetes, nearly all respondents drew on their experience with an existing diabetes improvement program in their health system. Briefly, this program alerts clinical staff to order tests/procedures if a patient has a diagnosis of diabetes in their medical record. Clinicians are then scored based on their adherence to diabetes quality measures for this population. In discussing the quality improvement program, it became apparent that clinicians are more conscientious about how they enter and manage the data in the EHR system with the knowledge that the data will be used in quality reporting. One clinician noted that there is more of an effort to obtain and document relevant lab values that might not normally make it into the record, such as lab results from institutions outside of the health system. He noted, “because of the quality program, we call (for the labs) and all that.” As described in Chapter 2, multiple respondents also noted that they are more careful to document a diabetes diagnosis on the problem list. One clinician explained, “With the quality program, usually there’s a pressure to put the diagnosis in and you think about it more.”

4.3.2 Care Process

In general, clinicians supported the use of EHR data for identifying patients with diabetes for quality reporting. However, clinicians did express concern about a lack of refinement in the EHR algorithms used to identify these patients. Specifically, clinicians felt that some patients with diabetes are inappropriately included in denominators when measuring quality of care. For example, a physician might be told to apply a quality standard, such as keeping an HbA1c level below 7.0, when that level is not appropriate to the specific patient. One clinician explained, “One thing we have not done, in any
kind of sophisticated way, is really think about how to really identify the population we want to treat this way and defining who really, genuinely qualifies. The country got excited about keeping HbA1c below 7. But we got data that said for older adults keeping HbA1c that low actually increases the risks associated with hypoglycemia more than having risks of slightly elevated numbers.”

In addition to inappropriately applying treatment guidelines to patients who should not qualify (e.g. elderly, terminally ill, etc), respondents reported that the lack of sophistication of the EHR algorithms impacts clinician time. Clinicians reported that identifying those patients for whom the quality measures do not apply is time consuming for nurses and doctors. Clinicians must go through the EHR manually to determine whether or not the quality standard should be applied to the patient in question. One clinician reported, “One of my colleagues says it takes him 5 to 6 minutes per patient to find out what’s happening. ‘This patient is 92. I don’t think a colonoscopy is necessary.’ And then you scroll down through the report and find that you discussed the colonoscopy and the patient doesn’t want it. So it is a tremendous amount of extra work for the physician with no additional help at all. “

Overall, clinicians emphasized the importance of treating the individual patient, rather than depending on an EHR-based algorithm. One physician summarized this sentiment saying, “I think algorithms can be used, but you need to be very conscientious that you keep the patient at the center of it and not the algorithm at the center of care and evaluation.”

4.3.3 Insurance Issues

The most frequently mentioned unintended consequence involved the issue of insurance coverage. Clinicians felt that, in the rare event that the EHR data generates an incorrect diagnosis and documents the diagnosis in the record, patients could be denied
insurance coverage or made to pay higher premiums. Respondents felt such misdiagnoses in the record could impact life insurance, long term care insurance, and health insurance coverage. As one respondent speculated, “The insurance company gets wind of that with the record and says, ‘Oh my god, you’re diabetic, we’re going to triple your premiums on your life insurance policy’ or ‘you’re not going to get health insurance because of this.’” In most cases, respondents were speaking hypothetically. However, a few clinicians told stories about insurance companies denying coverage based on incorrect diagnoses in the medical record, highlighting the importance of accurate diagnosis information in the record.

4.3.4 Patient Anxiety

Clinicians also expressed concerns about a patient learning of their diagnosis before a physician has the opportunity to discuss the diagnosis, if the diagnosis derived from the EHR algorithm is automatically inserted into the record. Clinicians agreed this scenario could result in unnecessary patient anxiety. Participants described three different ways a patient could learn of a diagnosis before the physician has notified the patient. First, clinicians explained that many of their patients access their EHR via a patient portal. They noted that if a computer algorithm adds a diagnosis to the problem list, there is potential for a patient to view a diagnosis via the patient portal before a clinician has the opportunity to contact the patient, particularly if the clinician is not made aware that the diagnosis has been added to the record. One clinician recalled a few instances when this scenario occurred with a chronic kidney disease (CKD) quality program, in which diagnosis of CKD is automatically added to the problem list based on lab measures obtained in the EHR system.

Second, clinicians reported there is also potential for a patient to be notified of
<table>
<thead>
<tr>
<th>Topic</th>
<th>Comments</th>
</tr>
</thead>
</table>
| Care Process  | So certain types of things we should make exceptions for. So something like obviously terminally ill people we would just try to keep their sugar in a range that they're not in the hospital. We don’t need to treat them down to 5.6 or 6.0 or whatever. If they’re 8, that’s ok.  
I think for some elderly frail people its too aggressive. They are too likely to get dizzy and fall and have other problems. And also, I should mention the A1c target of 7 is too aggressive for a certain number of diabetics.  
Patients are individual and not everybody fits into a nice, tidy box… Yes, I mean some patients can't tolerate a blood sugar under 160, so you treat the individual.  
But at the same time, with data mining, you are just looking at numbers and there are lots of other things you have to consider. There are a lot of serious implications.  
Say there’s a dialysis patient, and they don’t make any urine and they are diabetic and their microalbumin alert flags. Well, they’re already on dialysis anyway. Even if you can make a little bit of urine, what’s the point?  
I mean, it’s frustrating at times and it’s frustrating with some of the very elderly. We probably shouldn’t be bringing their blood pressures down to what our goal is for all our diabetics right now, less than 130 over less than 80. |
| Insurance     | I had somebody that got denied for long term care insurance… And they used an encounter diagnoses from 2007 that were used to order some vascular studies. You know, he comes in with this letter and says, ‘I don’t have this.” I couldn’t find it anywhere. I ended up doing a search through all the encounters and eventually found it. So, some of these insurance companies go through those records with a fine-toothed comb.  
I guess in terms of life insurance and health insurance for people, you know, now they have this label put on them.  
So, if you use it as a criteria for putting a name of a person and a diagnosis together for insurance purposes, so that their premiums are high because of that, I don’t think that’s right. |
| Patient Anxiety | That’s huge because now, with patients having access to the problem list because they have access through the patient portal, and they see that, forget about it. I don’t want it creating anxiety for them, which, in some cases, chronic kidney disease has. It creates a lot of unnecessary anxiety.  
Well patients not being aware of it and suddenly becoming aware of it. If they’ve got access to the patient portal and all the sudden it says they’ve got diabetes, you know, ‘I didn’t know I had diabetes.’  
You’re going to be on that list and nobody’s contacted the patient, so that would be an issue.  
No patient should actually get a letter without the physician knowing about it. And the letter is signed “your physician” or even the physician’s name on it. And the patient is saying, “You sent me the letter.” And there’s nothing in EPIC that tells us a letter was sent. And it makes a very complex discussion with the patient. |
their diagnosis via a letter, before the physician has the opportunity to discuss the
diagnosis. Clinicians described a quality initiative in their health system in which a letter
is automatically sent to patients regarding a diagnosis of CKD when an EHR
algorithm automatically adds the diagnosis to the problem list. Third, one physician felt
that it was possible that a patient could discover a diagnosis after a medical record
review by an insurance company. The respondent explained, “There have been times
where people had a diagnosis and maybe were denied insurance, especially life
insurance. So they go to buy a house, and they need to have life insurance to cover the
mortgage, and they end up getting denied because they have diabetes and they didn’t
fully realize that or whatever.”

The unintended consequences are more severe when the EHR-generated
diagnosis is incorrect. Two respondents told stories of patients who incorrectly received
automated letters notifying them that they had CKD. In both cases, the patients did not
have CKD and the physicians were not told the letters were being sent. One physician
described the fall-out: “It was a mess that took 6 months to clean up. Doctors were
accused of neglecting patients. Yeah, so I think we need to be very cautious. One day
we are going to get sued.”

4.4 Discussion

Overall, clinicians endorsed the use of EHR-data to identify patients for quality
measures. However, clinicians had concerns about the potential for unintended
consequences, ranging from patient anxiety to denial of health insurance. In many
cases, clinicians went beyond describing hypothetical scenarios, and provided specific
examples of unintended consequences. Despite their concern regarding the negative
consequences, clinicians also highlighted an important positive unintended consequence
of using EHR data for quality reporting.
The use of EHR data for quality measures is not unique in terms of its potential for unintended consequences. There is evidence in the literature of the unintended consequences across a range of HIT applications.\textsuperscript{47-52,96,97} Frequently, HIT applications, such as computerized physician order entry (CPOE), have been developed to improve safety and quality, but can result in negative, unintended consequences to the patient, physician, and/or health system.\textsuperscript{50,98} For example, Han and colleagues reported an unexpected increase in mortality with the implementation of CPOE, due to the unanticipated delays in administration of critical medication.\textsuperscript{52} Other studies found that HIT innovations eliminated some types of medical errors, but created new, sometimes more serious errors.\textsuperscript{54,96,97} We found that using EHR data to diagnose patients for quality purposes has the potential to cause unnecessary patient anxiety, impact patient access to health insurance, and result in inappropriate application of general treatment guidelines. At the same time, our study revealed the potential for improved documentation in the EHR.

The clinicians in our study expressed concern about the inability of EHR algorithms to account for patients who should be excluded from certain quality measures. This finding is consistent with studies that have revealed the inability of EHR algorithms to effectively account for medical exceptions. Persell et al. report that allowable exceptions for quality measure failures for coronary artery disease (CAD) are not reliably captured by automated searches of the EHR.\textsuperscript{99} In another study, Baker and colleagues examined the ability of automated EHR data to measure quality of care of outpatients with heart failure and found that EHR algorithms missed many exclusion criteria for medications documented in the providers’ notes.\textsuperscript{101} The consequence of missing these exclusions include underestimates of quality scores and pressure to meet quality guidelines even for patients who may not benefit, or could suffer, from these
guidelines. Wachter describes scenarios where clinicians were persuaded to modify their care delivery to be compliant with quality measures, regardless of their appropriateness to the specific patient. One option for addressing the challenge of accounting for exceptions is to give clinicians the option to record when a patient should be considered an exception to the quality measure. Documentation of exceptions would result in more accurate quality assessments and enhance clinical care by allowing the physicians to track the reason a patient did not receive the recommended treatment.

Clinicians were most concerned about the negative consequences of quality programs that automatically document a diagnosis in the EHR based on an EHR algorithm (e.g. automatically add a diagnosis to the problem list). One potential consequence discussed by a few respondents was the potential for a patient to see the diagnosis via a patient portal before the doctor has the opportunity to discuss the diagnosis. Patient access to the EHR data is not unique to the institution we studied. Healthcare systems such as Kaiser Permanente and the Veterans Administration provide patients access to their EHR via patient portals and the problem list is frequently among the data elements to which patients have access. If healthcare systems are going to implement quality programs that result in automatically adding EHR-derived diagnosis to the record, healthcare systems should modify workflows to ensure that patients are not notified about a diagnosis before they have the opportunity to discuss the diagnosis with their clinician.

Documenting an EHR-based diagnosis also has the potential to impact insurance coverage. Clinicians feared that diagnoses that are automatically added to the medical record, as opposed to having them stored in a separate quality database or registry, could result in higher insurance premiums or denial of coverage for affected patients. While insurance companies cannot access medical record data without signed
permission from the patient, most companies require that patients complete a form authorizing the company’s access to the patient records. A misdiagnosis could result in time-consuming attempts on the part of the patient and physician to correct the medical record and end with a long appeal process. Health systems that use EHR algorithms to generate diagnoses should give careful consideration as to whether the diagnosis needs to be added to the medical record or whether the diagnosis could be stored outside the record and used only for the quality programs for which it was created. Health systems should also consider requiring a physician review of automatically generated diagnoses before the diagnosis is added to the record, a strategy that has had success in some health systems.

Not all of the unintended consequences we identified were negative. Clinicians noted that the knowledge that their quality was being measured using EHR data motivated them to improve their documentation in the record. This is an important finding, given that many of the promised benefits of EHRs, such as improved quality of care and reductions in medical errors, are dependent upon complete and accurate documentation. Missing data is a problem in primary care practices. Smith and colleagues found that clinicians report they are missing important data in one out of every seven patients visits. Documentation continues to suffer with the transition from paper to electronic records. Hann and colleagues reported that documentation of basic health history and preventive indicators in primary care setting with an electronic system is no better than documentation using a paper system. If EHR-based quality improvement programs drive clinicians to improve their documentation, it could, not only make the quality reports more accurate, but could result in more effective EHR-based innovations in general.

4.4.1 Limitations
We acknowledge two limitations of our study. First, this study, like many studies evaluating HIT, was conducted in a health system that is considered one of the early adopters of HIT. As a result, we likely did not uncover some of the unintended consequences of using EHR data to diagnose a patient in a system with less experience in using EHR data. It is likely that the health system we studied has more complete EHR data and more experience creating algorithms for diagnosing patients than health systems with more recent EHR adoption. However, we wanted to identify the unintended consequences of this HIT application, independent of general data issues that can occur with a new information system. We felt that selecting a health system with more experience with the EHR would allow us to achieve this goal.

Second, as previously noted in Chapter 2, the clinicians who agreed to participate in our study may have been more interested and knowledgeable about the use of EHR data for quality reporting than those who did not agree to participate. As a result, these clinicians may be more aware of the potential consequences of applying EHR data to quality programs than clinicians in general. We took several steps to minimize the potential for selectively recruiting these types of clinicians. First, we requested that the administration provide us with a recruitment list of clinicians with a range of knowledge and experience in EHR data and quality reporting. Second, we provided a financial incentive to participants with the goal of attracting physicians regardless of their interest in the study topic. Given our response rate of 46%, relatively high for clinicians, we believe we effectively minimized this potential source of bias.

4.5 Conclusion

The federal government is providing incentives for health systems to use their EHR data for quality reporting. In addition, federal incentives for EHR adoption will likely result in an influx of new EHR adopters. As a result, the use of EHR data for quality
reporting is likely to expand across health systems with a range of expertise in EHR data. Future research should explore the unintended consequences of applying EHR data to quality measures at other institutions, particularly among new adopters who might be presented with a new set of unintended consequences. Researchers should investigate methods of minimizing negative consequences, capitalize on positive consequences, and manage the inevitable unintended consequences of using EHR data for quality reporting.
Chapter 5
Discussion of Findings: Implications and Future Research

The previous three chapters describe three main findings of a project designed to explore the use of EHR data for performance measurement in diabetes care. This chapter highlights the key findings of the project, discusses the implications of these findings, and suggests topics for future research.

5.1 Review of Findings

In Chapter 2 we reported our findings from interviews with clinicians regarding EHR data entry practices and the potential impact of these practices on performance measurement. Clinicians indicated that a diagnosis on the problem list would be the most sensitive and specific method of identifying patients with diabetes. Clinicians also endorsed the use of an encounter diagnosis of diabetes, but noted that the use of this data field could result in some disease misclassification. Specifically, when an encounter diagnosis is used as a rule-out diagnosis (e.g., when diabetes is suspected, but not confirmed), a patient without diabetes could be incorrectly included in the target population. Alternatively, if a patient with diabetes fails to come in to the office for treatment, that patient will not have an encounter diagnosis of diabetes and will be incorrectly excluded from the target population.

When providers made their recommendations for methods to identify patients with diabetes, they took into consideration factors that influence their data entry practices.
During the interviews, it emerged that organizational factors have influence over data entry practices. We identified three organizational factors that play a role in EHR documentation: quality improvement initiatives, organizational pressure, and workflow.

Quality improvement initiatives and organizational pressure from the health system were deemed responsible for the providers’ attentiveness to keeping the problem list accurate and up-to-date. Respondents were motivated to keep the problem list up-to-date to ensure they received diabetes best practice alerts. In addition, organizational leaders pressured clinicians to keep accurate problem lists by communicating the importance of the problem list and monitoring clinician problem list activities. As one respondent explained, “I think it’s one of the things that the system has made a priority so you’re more focused on things that are out there in front where you are getting measured and looked at.”

The third organizational factor, workflow, impacted the encounter diagnosis of diabetes, but not the problem list diagnosis. Providers explained that the EHR system in their organization allows for manual and automated entry of an encounter diagnosis of diabetes. Specifically, in addition to entering the diagnosis manually, there are systems in place that automatically generate an encounter diagnosis if a diabetes-related order (e.g. order for HbA1c lab) is placed. As a result, patients with a diabetes-related order are more likely to have a diabetes encounter diagnosis than patients who do not have a diabetes-related order.

As a complement to the information gained in the provider interviews, we conducted a quantitative study of EHR data to evaluate different methods of identifying patients with diabetes for diabetes performance measures (Table 3.1). We measured the sensitivity and specificity of eight different EHR-based definitions to identify primary care patients with diabetes. Next, we calculated the proportion of patients identified by each
method who met diabetes quality standards for blood pressure, LDL, and HbA1c control. Finally, we determined whether these methods selectively identified patients who were more or less likely to meet these quality standards. Much of what we learned from our clinician interviews helped to explain our quantitative findings. However, findings were inconsistent in some areas.

All of the definitions we evaluated had high specificity (>99.5%) and moderate to high sensitivity (65 to 100%). Definitions that classified patients as having diabetes if they had at least two encounter diagnoses of diabetes were the only definitions to have less than 100% specificity. These definitions misclassified two patients without diabetes patients as having diabetes. A detailed review of the two patient records revealed that the patients had suspected diabetes, but were ultimately ruled not diabetic. This finding is consistent with clinician reports stating that an encounter diagnosis is sometimes used as a rule-out diagnosis.

The definition that required the presence of a diabetes diagnosis on the problem list (definition 2) had lower sensitivity (75%) than all other definitions except for the definition that required two elevated HbA1c measures (definition 3). This finding is inconsistent with the reports from clinicians that the problem list would be the best way to identify patients with diabetes. There are at least two explanations for this incontinency. First, adherence to problem list documentation may be higher among study participants than in other primary care providers in the Geisinger Health System. Alternatively, respondents may have overestimated the accuracy of the problem lists. Given our efforts to reduce selection bias in our recruitment strategy, we do not believe that the study participants use the problem list differently from other clinicians. The literature contains a number of examples of clinicians overestimating their
We suspect that clinicians in our study believed they and/or their colleagues maintain more accurate problem lists than they actually do.

The definition used to identify patients with diabetes did not appear to have a large impact on the quality performance for blood pressure, LDL, or HbA1c measures. Patients identified across the eight definitions scored within 1.7 percentage points of each other on the blood pressure measure, within 4.1 percentage points of each other on the LDL measure, and within 2.2 percentage points of each other on the HbA1c measure. However, we determined that the EHR criteria used in each of the definitions did selectively identify patients who had better quality performance scores. Specifically, definitions that classified patients as having diabetes if they had at least two encounter diagnoses associated with diabetes (definition 1, 5, 6, and 7) selected patients with better quality performance, and excluded patients with poorer performance.

While the original intent of our clinician interviews was to learn about the validity of using the EHR for performance measures, during these interviews it emerged that there are unintended consequences associated with using EHR data to measure performance. As a result, we pursued a line of inquiry that increased our knowledge about these unintended consequences. Clinicians summarized three categories of negative unintended consequences associated with using EHR data for diabetes quality initiatives. First, EHR-based quality measures are frequently not sophisticated enough to identify patients who should be excluded from performance measurement, including the elderly or terminally ill patients. This is an important issue for at least two reasons. First, failing to account for exceptions could result in underestimating the performance of a physician who did not meet quality standards because patients should have been excluded from the performance measure. Second, and more importantly, clinicians may
incorrectly apply quality standards to patients who should be excluded, to the detriment of the patients’ health.\textsuperscript{103}

The other negative consequences involved the consequences of having a diagnosis generated by an EHR algorithm and having that diagnosis automatically entered into the EHR. First, clinicians feared that the algorithm could result in a misdiagnosis in the record that would leave patients vulnerable to insurance denials. Second, clinicians feared that diagnoses automatically generated and added to the medical record could cause anxiety among patients if patients learned of the diagnoses before physicians had the opportunity to notify the patients. In both cases, clinicians were able to provide specific examples of these types of occurrences.

Clinicians were also able to highlight a positive unintended consequence of using EHR data for quality programs. As discussed earlier, the diabetes quality program that was driven by a problem list diagnosis of diabetes acted as a motivator for physicians to keep the problem list up-to-date. While our analysis of the EHR data leads us to conclude that the clinicians in this study overestimated their use of the problem list, the actual use of the problem list (sensitivity: 75\%) appears to be better than reported elsewhere in the literature.\textsuperscript{73-77} In part, the quality improvement program may be responsible for the relatively better use of the problem list.

5.2 Implications

The findings from this study provide important information to administrators of quality reporting programs; health system leaders participating in EHR-based performance measurement; and patients, providers and payers viewing performance data. Moreover, these findings have implications beyond performance measurement. Health services researchers and epidemiologists using EHR data can benefit from a
better understanding of EHR data and its use in identifying disease populations. The implications of the study’s findings are detailed below.

One of the key findings of this project is that the method used to identify patients with diabetes in the EHR can be associated with diabetes care performance. Specifically, we found that the use of certain criteria (e.g., 2+ encounter diagnoses) excluded patients from the target population who performed poorly on performance measures. To date, only one other study has evaluated EHR-based diabetes performance measures for selection bias, and this study looked only at encounter diagnoses counts and was confined to Medicare beneficiaries. Our new finding has important implications for performance measurement, particularly for performance measurement used in public reporting of health performance. We describe these implications below.

First, when comparing health systems on performance in diabetes care, it is essential that participating providers use the same method for identifying diabetes patients. When health systems are unable to apply the same methods, for example, when systems have access to different types of EHR data, comparisons should be made with extreme caution. Second, when developing performance measures, administrators of quality reporting programs must provide detailed information as to how to identify patients with diabetes. For example, administrators cannot simply say, “Include all patients with a diabetes diagnosis.” This example of loose criteria could lead one organization to use an encounter diagnosis while another may use a problem list diagnosis, thus leaving the measures vulnerable to manipulation. Finally, when a provider or health care organization reports performance measures, the health care provider should be clear when describing the methods used to identify the target population. A review of current public reporting programs indicate that reports currently
lack the level of detail described above to allow the public, payers, or health systems to interpret reports or make fair comparisons across organizations.\textsuperscript{7}

Another key finding of our study was learning about the influence that organizational factors have on EHR data entry and the impact this can have on performance measures. This finding is essential to consider as organizations such as CMS move towards standardizing EHR-based performance metrics.\textsuperscript{114} Whenever possible, organizations developing these standards should attempt to use EHR data elements that are not sensitive to organizational influence. Given that it is unlikely that many data elements will be completely free from the influence of organizational factors, comparisons of performance across organizations should, once again, be made with caution.

The relevance of these findings goes beyond performance measurement. The EHR is being increasingly used for secondary purposes such as clinical trials, comparative effectiveness research (CER), and general epidemiological and health services research. The Federal Coordinating Council for Comparative Effectiveness Research acknowledged in its 2009 annual report that the success of CER is largely dependent on health information technology.\textsuperscript{115} As with performance measures, CER requires the identification of a target population to answer research questions. If treatments for diabetes are compared in a population that excludes diabetics with poorer cholesterol or blood pressure control, results of the research cannot necessarily be generalized to that excluded group. Researchers should consider the findings in this report when using EHR data to identify patients for research.\textsuperscript{116}

The influence of organizational factors on maintenance of the problem list has important implications for health systems that are striving to improve the use of the problem list in the EHR. The problem list is considered a key aspect of the EHR. It helps
practitioners identify the most important health factors for each patient, allowing for customized care. The federal EHR incentive effort recognizes the importance of the problem list, requiring that professional and hospitals participating in the meaningful use program must code problem lists.\textsuperscript{117} Other organizations can learn from Geisinger's efforts to improve problem list documentation.

This is the first study to report on the unintended consequences of the application of EHR data in quality programs. EHR adoption will likely increase with federal incentives for adoption and our findings are critical for organizations to consider as they begin to leverage their EHR data for quality programs. With better understanding of these consequences, organizations can implement strategies to avoid these negative unintended consequences. First, when using EHR algorithms to generate diagnoses for performance measures, these diagnoses should be stored outside the electronic record until the patient's physician can confirm the diagnosis. This strategy should minimize the potential for the presence of a misdiagnosis in the record, and also reduce the chance a patient will discover a diagnosis before a physician had the opportunity to notify the patient. Second, when building performance measures into the EHR, health systems should consider building a data field into the EHR that allows clinicians to document when patients should be exceptions to the quality standard. Documentation of exceptions would result in more accurate quality assessments and enhance clinical care by allowing the physicians to track the reason a patient did not receive the recommended treatment.\textsuperscript{101}

5.3 Future Directions for Research

This project was an in-depth study of the use of EHR data for identifying patients with diabetes for diabetes performance measures in a single health care system. This project was not intended to make a final recommendation about the EHR algorithm to
use to identify the target population for diabetes performance measures. Further, the findings of this project should not be used to endorse or refute the use of EHR data for performance measurement. Instead, the results of this report should be used as guidance when developing diabetes performance measures and creating quality reporting programs. The results of this study point to directions for future research in the area of diabetes performance measures and the use of EHR data in performance measurement in general.

This study determined that including a patient with a problem list diagnosis, ≥2 encounter diagnoses, a diabetes medication order, or ≥2 elevated HbA1cs is the most sensitive method to identify patients with diabetes in the Geisinger EHR. However, as we learned in this study, organizational factors play a role in EHR data entry. Therefore, this diabetes definition is not the final word as to the EHR criteria that should be used for identifying patients with diabetes. The eight EHR definitions we examined, and perhaps other definitions, should be studied in other health systems. It is particularly important that this research is expanded to institutions likely to have organizational influences that differ from GHS. For example, future research should target institutions with a different physician employment structure, less mature EHR systems, or fewer EHR-based quality programs already in place. Expanding this research to other organizations is a critical step toward defining a standard method of identifying patients with diabetes for performance measures, as well as for epidemiological and health services research.

Just as developing a standard method to identify the target population for performance measures is important, it is critical to develop a standard numerator definition (i.e., criteria for defining who meets the quality standard). For example, there could be multiple ways to determine whether a patient meets the criteria for good blood pressure control (e.g., Use last recorded blood pressure, use average of last three blood
pressure measures). Studies have reported that, depending on the EHR criteria used, automated EHR reports can mislabel patients as meeting or not meeting quality standards.99 Evaluating numerator definitions was outside the scope of this project. However, in moving towards standardized measures, researchers will have to evaluate this aspect of performance measures as well.

While this study is the first to identify unintended consequences of using EHR data for performance measures, it is unlikely that we identified all of the potential consequences. EHR-based performance measures in health systems with less developed EHRs may be subject to another set of consequences that we did not uncover. For instance, missing data present in less advanced EHRs may account for a number of consequences we did not identify. As more health systems start to participate in EHR-based performance programs, it is essential that health system leaders assess these programs for unintended consequences and share their findings with other health care organizations. As was the case in our study, many unintended consequences can be mitigated if organizations are aware of the potential for these consequences when developing their performance programs.

This study explored the use of EHR data alone for developing quality measures in diabetes. However, despite the documented limitations of claims data and the challenges we uncovered associated with EHR data, there is a movement to use the combination of information offered by both data sources for measuring performance. The Agency for Healthcare Research and Quality (AHRQ) recently sponsored several research and pilot projects to explore methods for using clinically enhanced claims data to improve performance measures for quality reporting.118 In addition, the National Quality Forum (NQF) is in the process of identifying ambulatory care measures suitable for public reporting and quality improvement programs that are derived primarily from
clinically enhanced administrative data. Future research should explore the advantages of using clinically-enhanced claims data, versus depending on EHR data alone for performance measurement.

Finally, the focus of this project was on performance measures in diabetes care. However, diabetes is just one of many conditions for which performance measures have been created. Research similar to the work described in this paper should be conducted across the spectrum of disease areas that are being measured. EHR data entry practices and the methods of identifying a disease population may differ based on such disease characteristics as chronicity and ease of diagnosis. Future research should expand to other disease areas, such as cardiovascular disease, that are frequently the subject of performance measurement.

5.4 Concluding Thoughts

There is much, and frequently well-deserved, enthusiasm for the potential for EHR systems to drive improvements in health care quality. One of the proposed paths towards this achievement in quality improvement is through the use of EHR data for performance measurement. However, in the rush to solve the quality problems of our health system, it is imperative that we take the time to fully understand EHR data and the consequences of its use in performance measurement. Only after we apply this knowledge can we fully reap the benefits of EHR technology.
References


76


48. Ash JS, Sittig DF, Dykstra R, Campbell E, Guappone K. The unintended consequences of computerized physician order entry: Findings from a mixed


81. Chen T, Chung K, Lin I, Lai M. The unintended consequence of diabetes mellitus pay-for-performance (P4P) program in Taiwan: Are patients with more comorbidities or more severe conditions likely to be excluded from the P4P program? Health Services Research. 2011; 46(1,p1): 47-60.


APPENDIX A: Clinician Interview Guide

INTRODUCTION AND BACKGROUND

First, let me thank you for agreeing to participate in our research project. The Center for Health Research is studying the use of the electronic health record systems (or EHRs) in measuring healthcare quality.

For this part of the study, we are interviewing physicians/clinicians to get a better understanding of how they enter data into the electronic health record. Specifically, we want to learn what fields in the electronic health record physicians use to document diagnoses and your reasons for using those fields. Your answers may help us determine what fields in the EHR should be used when measuring quality. For example, your responses may help us decide the best way to identify patients with diabetes for quality reporting in diabetes care.

For the purpose of this interview, we are particularly interested in how you use the electronic health record when diagnosing and managing patients with diabetes. We selected diabetes because it is one of the diseases for which there is a number of existing quality measures, including measures that permit quality reporting from the EHR.

This interview will last about 30 minutes. During this interview I am going to ask you some open-ended questions about your experience with the EHR, and about how you use the EHR to record patient diagnoses. In addition, I am going to ask you some questions about the different reasons you document diagnoses in one part of the EHR versus another. With your permission, we would like to record the interview so that we can more completely and accurately capture your comments. If at any time you wish to make an “off the record” comment, please let us know and we will stop the tape at your request and restart it when you are ready to continue. (Start recording only after reading “Voluntary and Confidential” section and obtaining verbal consent to record- See next page.)

VOLUNTARY AND CONFIDENTIAL

Before we begin this interview, let me make sure that you understand that:

Taking part in this research study is voluntary. You may choose not to take part or may leave the study at any time. If you choose to participate, you can answer or not answer any questions, as you wish. Your decision to participate or not participate in this research project will not affect in any way your job or position.
We consider this discussion to be confidential. Your participation is confidential in the sense that your name will not be used in any reports or articles.

Efforts will be made to keep your responses confidential. We have attempted to minimize the potential for a breach of confidentiality by training study personnel in human subjects’ protection and by following Institutional Review Board (IRB)-approved procedures for protecting confidentiality and maintaining data security. If data or information from the research study is submitted for publication in a medical journal, or is presented at a medical meeting, your identity as a research participant will not be revealed. Access to interview transcripts or data will be restricted to the research team. Any transfer of electronic files will be in a password-protected format. Electronic and hard copies of data will be stored without personal identifiers. We cannot guarantee absolute confidentiality. Your personal information may be disclosed if required by law. Geisinger’s Institutional Review Board may need to have access to your research records for quality assurance.

For questions, concerns, complaints, or if you feel you have been harmed as a result of study participation you may contact Ms. Annemarie Hirsch at 570-214-4880.

For questions about your rights as a participant in this study, or to discuss other study-related concerns or complaints with someone who is not part of the research team, please contact the Human Research Protection Program staff of the Geisinger Institutional Review Board (which is a group of people who review the research to protect your rights) at (570) 271-8663.

Do you have any questions about this interview or the study?

Would you be interested in participating in this study?

If no → OK. Thank you very much for your time today. Goodbye.
If yes → continue.

Great! Thanks for agreeing to participate in this study. Is now still a good time for you to speak or would you like to schedule an appointment for some time in the future?

Do you mind if we record the interview? [If yes, “Okay, I’m starting the recorder.” If no, “Okay, I will just take notes on our discussion.”]

OVERVIEW OF INTERVIEW TOPICS

As I mentioned briefly, during this interview, I will ask you a series of open-ended questions about a variety of topics to get your perspectives. Our questions are grouped into four general topics:
Section 1: History and Background
Section 2: Documentation in EHR
Section 3: Motivation for Documentation in EHR
Section 4: Quality measures and EHR

Some of my questions will be general and others will be specific to the management of patients with diabetes. Do you have any questions before we begin?

BEGIN INTERVIEW

Section 1: HISTORY AND BACKGROUND: First I am going to ask you some general questions about your role at Geisinger and your patient population.

1) To start, could you please tell me about your practice and your role at Geisinger?
   a) Probe: How long have you been in this role?
   b) Probe: How long have you been at Geisinger?
   c) Probe: How long have you been practicing?

2) Could you give me a sense of the size and scope of your practice/group/department?
   a) Probe: How many patients do you typically see? (per week, per month?)
   b) Probe:
      i) Are patients typically assigned one physician/clinician in the practice as opposed to having their health managed across multiple providers in the practice?
      ii) How often do you treat patients primarily managed by other physicians in your practice, due to scheduling issues, etc?

3) About how many patients with diabetes do you see per month?
   a) Probe: How many times a month do you see diabetes patients whose diabetes is primarily managed by other physicians in your group? In Geisinger?
   b) Probe: How many times a month do you see diabetes patients whose diabetes is primarily managed by physician outside of the Geisinger system?

4) Does your department participate in any quality improvement programs? Can you describe them?
   a) Probe: Do you or your department receive incentives for meeting certain performance goals? If so, describe.
   b) Probe: Does your department participate in any quality improvement programs in diabetes management?

Section 2: DOCUMENTATION IN EHR: The next set of questions is about how you use the EHR to document diagnoses, specifically diagnoses of diabetes.

5) When you think of the range of fields you use in the EHR, where do you record diabetes diagnoses in the EHR?
   a) Problem list? (possibly jump to questions below on problem list)
b) Encounter diagnoses? (possibly jump to questions below on encounters)
c) Reason for visit?
d) With medication order? (possibly jump to questions below on meds)
e) With lab order? (possibly jump to questions below on lab orders)
f) In progress note?
g) When you record a diagnoses in one location in the EHR, does it automatically populate another location in the EHR?

6) Do you record diagnoses of other diseases in the same way? (e.g. more acute diagnoses). Why/Why not?

7) When treating a patient, where do you usually look in the record to determine what diagnoses they already have? (i.e., diagnosis previously documented by you or another physician)
a) Why? Ease of checking that location; more reliable field; etc?

8) I understand that some of your patients may be managed by multiple physicians. When reviewing the medical records of these patients, how much variation do you see in how/where physicians document diagnoses in the EHR?
a) Does it depend on the condition? Do you find physicians are more or less consistent when it comes to documenting a diabetes diagnosis?
b) Do most physicians use the problem list in the same way?
c) Do you know where to look to find diagnosis information? Does it vary from physician to physician?
d) Does it vary more or less with a diagnosis such as diabetes?

Section 3: MOTIVATION FOR DOCUMENTATION- For this next part of the interview, I am going to ask you some follow-up questions about how and when you use specific fields in the EHR. For each field in the EHR, I am going to ask you some general questions, and then I am going to ask how you use the field for patients with diabetes.

9) I would like to learn more about when a diagnosis is listed as an encounter diagnosis.
a) Is this a required field for every encounter?
b) Do you list a diagnosis as an encounter diagnosis if the diagnosis is the" reason for visit?"
c) Do you list a diagnosis because you are planning to prescribe a medication or order a procedure?
d) Do you consider insurance coverage?
e) Do you consider reimbursement?
f) How do you decide what to list as a primary versus a secondary diagnosis?

10) When treating a patient with diabetes, when do you list diabetes as the encounter diagnosis? For example, if a diabetic comes to the office for a matter unrelated to diabetes, do you still list diabetes as an encounter diagnoses?

11) I would like to learn more about the problem list. When would you document a diagnosis in the problem list?
a) Are you required to use the problem list?

b) Do you use the problem list for acute illnesses, chronic illnesses, or both?

c) Does your organization provide any incentives for keeping the problem list up-to-date?

d) Do you list a diagnosis because you are planning to prescribe a medication or order a procedure?

e) Do you consider insurance coverage?

f) Do you consider reimbursement?

g) What type of diagnosis would you NOT document in the problem list? Why?
   i) Acute illness?
   ii) Chronic illness?

12) When in the course of treating a patient with diabetes do you typically add diabetes to the problem list?
   a) Do you add when diabetes is strongly suspected?
   b) When you treat a patient who has clearly had diabetes treated at Geisinger for many years, do you confirm that diabetes is on the problem list?

13) In general, do you ever remove problems from the problem list or indicate they are “resolved”?
   a) If yes, why? (no longer active, misdiagnosis, etc)
   b) Do you delete problems from the problem list that have been added by another provider?
   c) Is there ever an event that a diagnosis of diabetes is changed from active to resolved?

14) In general, when do you link a medication to a specific diagnosis?
   a) Is this a required field for you to complete a med order?
   b) How do you decide what to list as a primary versus a secondary diagnosis?
   c) Do you consider insurance coverage?
   d) Do you consider reimbursement?
   e) Do you consider any other factors when choosing to link a medication to a diagnosis?

15) When treating a patient with diabetes, when do you list diabetes as the medication diagnosis? For example, if you are writing a prescription for a condition completely unrelated to diabetes, do you still list diabetes as a indication for the medication?

16) Similarly, when do you link a procedure or lab order to a specific diagnosis?
   a) Is this a required field for you to complete?
   b) How do you decide what to list as a primary versus a secondary diagnosis?
   c) Do you consider insurance coverage?
   d) Do you consider reimbursement?
   e) Do you consider any other factors when choosing to link a procedure/lab to a diagnosis?

17) When treating a patient with diabetes, when do you list diabetes as the procedure diagnosis? For example, if you are writing an order for a condition completely unrelated to diabetes, do you still list diabetes as a indication for the procedure?

Section 4: NEW SECTION: WHAT TO USE FOR QUALITY MEASURES

18) CMS and other organizations are using EHR for quality reporting. They use different strategies for identifying patients. What do you think is the best way to use the EHR to find all the patients with diabetes?
a) Problem list: why/ why not? – will people be missed, what %, who is typically not on the problem list? (older/younger; newly diagnosed)
b) Reason for visit: why/why not?
d) Prescriptions such as insulin: why/why not? (other conditions?)
e) Labs such as elevated HbA1c levels: why/why not?, what level, more than one? (other conditions, bad results?)

19) CMS is measuring management of LDL, HTN, and HbA1c in patients with diabetes. They are permitting EHR. Advantages/disadvantages to using EHR data?
   a) Use last value
   b) Use average value

20) Do you think using EHR data for measuring quality is a valid way of measuring quality in diabetes management?
   i) Advantages/Disadvantages

Section 5: FINAL COMMENTS

• That was the last formal question. Is there anything we have NOT talked about, that you believe is important to the issue of documenting in the EHR?

NOTE: After the first 3 interviews, the following question was added regarding unintended consequences:

“Do you have any concerns about the unintended consequences of using EHR data to identify patients for quality measures?”

THANK YOU!! so much for your time and participation. Your comments were extremely helpful.