Grant Final Report
Grant ID: R18HS17018-02

Medication Monitoring for Vulnerable Populations via IT

Inclusive Dates: 09/21/07 – 05/31/10

Principal Investigator:
Dr. Christoph Ulrich Lehmann

Team Members:

Sean Berenholtz [*]	Jill Marsteller [*]
David Bundy [*]	Marlene Miller [*]
Lilly Engineer [*]	Bryan Sexton [*]
Christen Fullwood [*]	Martine Uveges [*]
Lorraine Kelly [*]	Albert Wu [*]
Richard Thompson [†]	Jing Tian [†]
Byron Baldwin [‡]	Erica Isles [‡]
Harry Caughey [‡]	David Silver [‡]
Donna Enright [‡]	Donna Walters [‡]
Lori Farren [‡]	Sue Weimer [‡]

[*] Johns Hopkins University
[†] Johns Hopkins University Biostatistics Consulting Center
[‡] Baltimore Medical System (BMS)

Performing Organization:
Johns Hopkins University

Abstract

Purpose: A cross-sectional demonstration of the ability of interoperable health information exchange and the electronic health record (EHR) to provide useful quality and safety measures.

Scope: Use EHRs to understand better medication monitoring as it pertains to quality and safety in the ambulatory setting, using a set of national consensus measures and identify the organizational culture issues linked to accomplishing this for the vulnerable populations served by two clinics of Baltimore Medical System Community Health Center which had implemented EHR over a year ago.

Methods: A cross sectional demonstration project of the equivalency and usefulness of IT-based quality measures fed back to the provider via monitoring bulletin. A single arm, pre/post intervention study was conducted using data from each round of bulletin distribution. 2,013 study eligible subjects were summarized in bivariate tests by overdue status using the $\chi2$ test (categorical variables) and ANOVA (continuous variables) followed by logistic regression analyses.

Results: 1) Automatic queries of EHRs to identify patients eligible for quality measures are feasible and potentially far superior to manual reviews of EHR data; 2) Only 1 to 5 providers described reports of quality measure data as useful. The percentage of patients receiving monitoring after appearing on the bulletin was not associated with provider attitudes about it's usefulness; 3) Individuals at risk for medication-related toxicity frequently go unmonitored. Provider-specific bulletins can enhance receipt of recommended monitoring among target patients; and 4) Low teamwork and safety climate was seen in 2007, and these scores were not significantly different in 2009.

Key Words: medication monitoring, ambulatory, drug monitoring, drug toxicity, electronic health records, process assessment, safety attitudes, culture

Final Report

Purpose

The overall goal of this project was a practice-based cross-sectional demonstration of the ability of interoperable health information exchange and the Certification Commission for Healthcare Information Technology (CCHIT)-certified Misys 8.0 electronic health record (EHR) to provide useful quality and safety measures for the vulnerable populations served by two clinics that are part of the Baltimore Medical System (BMS) Community Health Center (CHC).

The quality measures being evaluated for use with an EHR had been developed by the National Committee for Quality Assurance (NCQA), supported by the National Quality Forum (NQF) for ambulatory care, and focus on the safety monitoring for chronic medications that are commonly used by patients with heart disease and diabetes mellitus.

The project's intervention included a monitoring bulletin or performance report card provided to physicians every two months to inform them of patients that require therapeutic monitoring tests for one or more of medication groups in the quality measures.

The project also evaluated the relationship between contextual factors (teamwork and safety climate at BMS) and provider assessments of EHR quality and safety data as useful and actionable as well as evaluating whether deployment of EHR quality and safety measurement efforts will improve teamwork and safety climate at the clinics. Thus, this project was designed to serve as a model for maximizing the utility of information technology in ambulatory settings among patients at high risk for substandard care.

Specific aims of the project:

- **Specific Aim #1**—Develop and implement via EHR accurate quality and safety measures focused on medication monitoring for vulnerable populations that are served by BMS CHC, and explore the factors that influence accuracy of EHR-derived measures.

  o **Hypothesis #1:** The positive predictive value for the EHR-derived quality and safety measures will be equivalent to that for chart review measures.

  o **Hypothesis #2:** The positive predictive value of EHR-derived quality and safety measures will not be influenced by patient medical and socio-demographic variables.

- **Specific Aim #2**—Develop and implement EHR-based quality and safety measures of medication monitoring for vulnerable populations that are served by BMS CHC that are useful to clinicians and senior leaders.

  o **Hypothesis:** 70% of frontline providers will describe reports of quality measure data as useful and 70% of clinic directors and BMS senior staff will describe rolled-up versions of reports of quality measure data as useful. "Useful" will be defined as

information that is delivered in a timely manner and be perceived to be acceptable, understandable, and actionable.

- **Specific Aim #3**—Develop and implement EHR-based quality and safety measures of medication monitoring for vulnerable populations that are served by BMS CHC that impact patient outcomes.

  o **Hypothesis:** When presented with feedback that patient is in need of an annual therapeutic monitoring test, providers will respond 100% of time.

- **Specific Aim #4**—Evaluate the relationship between contextual factors (teamwork and safety climate at BMS CHC) and provider assessments of EHR quality and safety data as useful (Specific Aim #2) and actionable (Specific Aim #3) as well as evaluate whether deployment of EHR quality and safety measurement efforts will improve teamwork and safety climate at CHC.

  o **Hypothesis:** Teamwork and Safety Climate scale scores of the Safety Attitudes Questionnaire-Ambulatory Version will be positively correlated with provider perceptions of usefulness of data and provider responsiveness to data and will increase by at least 10 points across 2 year time span of study.

Scope

Background

Since the Institute of Medicine (IOM) reports *To Err Is Human* and *Crossing the Quality Chasm: A New Health System for the 21st Century* directed a national spotlight on quality and patient safety a decade ago, most demonstration projects evaluating measurement of healthcare services, their related clinical outcomes, and quality incentives have been focused on non-ambulatory care settings. More recently, a much needed proliferation of national efforts focused on how to most effectively and efficiently improve performance measurement, data aggregation, and reporting in the ambulatory care setting by entities like the Ambulatory Quality Alliance & the Ambulatory Care Project of the National Quality Forum has been seen.

The recent IOM report entitled *Preventing Medication Errors* highlighted that across all settings of care including ambulatory care, quality lapses and errors related to medications are some of the most prevalent risks and the morbidity due to medication errors is costly. The high risk combination of medication and ambulatory care has a paucity of knowledge on epidemiology, risks, performance measurement, and solutions, as is evident from limited number of recent studies, most of which were also focused on immunizations. Of the phases of medication use, encompassed by prescribing, dispensing, administering, and monitoring, the monitoring phase has by far the least data particularly for the ambulatory care setting.

Thus, Medication monitoring is the least well understood phase of medication use as it pertains to quality and safety especially in the ambulatory setting. Data that is available highlights significant potential for patient safety improvement by improving rates of medication

therapeutic monitoring. It is not known how best to identify patients in need of medication therapeutic monitoring in the ambulatory setting although the promise of EHRs is clear. It is also not known if providers reliably understand how to link this process of therapeutic monitoring to differences in patient outcomes given the lack of national guidelines on laboratory monitoring, With the recent advent of a set of national consensus measures on medication therapeutic monitoring in the ambulatory setting, the time is ripe to explore the utility of using EHRs to accomplish this and identify the organizational culture issues linked to accomplishing this successfully. Such an exploration must include an analysis of the cost and efforts required to develop EHR based medication monitoring.

Context

Baltimore Medical System in Maryland started rolling out an EHR in all its practices about 2 years prior to the project start date. This permitted the effort to link the Johns Hopkins labs with the BMS EHR and opened several opportunities for providing electronic reminders to providers that were linked to laboratory tests. Prior to the study, BMS ordered labs on paper and then results we scanned into the EHR. Because the results were scanned, they were potentially accessible to the provider (many clicks to see all relevant labs) but could not be used to generate automatic reports and reminders. As part of the project, a major undertaking was the development of a bi-directional interface from Johns Hopkins Pathology Systems to BMS' EHR. As a result, laboratory results were received in the EHR as metadata that could be parsed and used for analysis.

Settings

Two federally qualified community health centers under the Baltimore Medical System, Maryland (BMS).

Participants

Our study population was the 15 providers at each of the Belair-Edison and Highlandtown clinic sites of BMS (both federally qualified community health centers), encompassing the medical specialties of Internal Medicine, Family Practice, and combined Internal Medicine/Pediatrics. These providers were typically between 30-60 years of age. During the course of the grant, there were some providers, who left BMS and whom we could not survey after their departure.

For each of the above mentioned providers, the patient population included all adult patients, who were prescribed one or more of the four medication classes of interest for annual monitoring in the previous calendar year as well as any BMS patients that newly acquired these diagnoses during the study period. The latter was done by using the pre-linked claims data based on diagnosis codes to select those patients with diagnoses of heart disease and/or diabetes mellitus since all of these patients are highly likely to be on at least one of the index medications, namely digoxin, statins, diuretics, and ACEIs/ARBs as specifically recommended by the NCQA at the time of study. We did not exclude any person based on race, ethnicity, or gender. Per pre-defined specifications of these measures, those patients with these diseases were included if they were 18

years or older and received at least a 180 day supply of at least one index medication the medications during the measurement year. Also per pre-defined measure specifications, any patients who were hospitalized during the measurement year were excluded since these patients may have received a monitoring event during the hospitalization. Since no metadata in the EHR directly identified hospitalizations, we used natural language parsing of EHR data to identify and exclude patients. Patients whose insurance stipulated that they receive laboratory studies from a laboratory not linked electronically to the centers' EHR were also excluded.

Incidence

One recent study estimated that adverse medication events are common among older persons in ambulatory care with an estimated incidence of more than 500,000 cases each year. Reaching into the ambulatory arena post hospital discharge, another study reported that 11% of patients discharged from the hospital experienced adverse drug events at home in the month after discharge related to the medications they were prescribed at discharge. Jha and colleagues reported that 2.3% of over 3,000 hospital admissions reviewed were caused by and adverse drug event while the patient was home.

Prevalence

Given the nature of these NQF Ambulatory Care Project medication monitoring measures, our grant predominantly focused on patients with heart disease and diabetes mellitus, both of which are high priority areas for quality work as identified by the Institute of Medicine. In 2006, the year prior to the study, the numbers of patients at the two clinic sites with heart disease and/or diabetes mellitus annually were estimated as 2423 (42%) for the Belair-Edison clinic site and 3406 (58%) for the Highlandtown clinic site with the total of 5829 (representing potential duplications based on having both conditions of heart disease and diabetes).

During the one year of monitoring bulletin distributions between August 2008 and August 2009, 2,013 patients were identified as having been prescribed target medications for at least 180 days in the preceding year and thus eligible for the study at both the clinic sites. Of these, 1,164/2,013 (58%) were never overdue for recommended laboratory monitoring during the study period. Also, at each cross-sectional measurement (bulletin round), 74-79% of eligible patients in the two centers were up-to-date for medication monitoring, with a slight downward trend over time. 58% of patients consistently received recommended monitoring during the study period.

Methods

The study design was a cross sectional demonstration project of the equivalency (compared to manual chart review) and usefulness of IT-based quality measures. The quality measures were deployed in the Belair-Edison and Highlandtown clinics that are part of an umbrella organization, the Baltimore Medical System (BMS) Community Health Center (CHC), comprising six FQHCs across inner-city Baltimore and Baltimore county serving a predominantly low-income, minority population. The main reason for choosing these 2 clinic sites was that these 2 sites had implemented the Misys EHR in mid-2006 and were robustly up on Misys EHR for over one year

by the start of this grant. The Laboratory studies, necessary to fulfill the aims of this study, were performed at a laboratory affiliated with a nearby tertiary academic hospital as part of usual care. The laboratory results were reported electronically to the study centers' electronic health record (EHR) via a bidirectional interface designed specifically for this study.

Specific Aim #1

This aim involves the development and implementation, via HER, of accurate quality and safety measures focused on medication monitoring for vulnerable populations that are served by BMS CHC, while exploring the factors that influence accuracy of EHR-derived measures.

Data Sources/collection. After a six month delay (on part of our sub-contractor) due to the complexity of developing and implementing the bi-directional interface required for this study, we developed and implemented a machine query of the EHR system Misys at Baltimore Medical Systems (Maryland, USA) of patients eligible for the 4 quality measures. Patients prescribed one or more of the target medications in the previous calendar year were included if they were 18 years or older, received at least a 180 day supply of the medication, and had not been hospitalized in the preceding year.

In order to allow the development of such a query, a Results-Only Interface was implemented (Results sent from the Lab system to Provider system) by end of 4th quarter of Year 1. Because we did not want to wait for another twelve month until we had enough laboratory data accumulated to identify patients overdue for laboratory testing, two years of Lab data history was back-loaded into patients' records via interface by end of end of 4th quarter of Year 1. The result reporting interface required to address the specific aims of the project continued to work well for the purpose of the project until the end of the data collection period of the project i.e., August 2009.

Method of Data Collection. EHR-based system generated patient flags from interface between BMS EHR Misys system and Johns Hopkins University laboratory data; chart review with standard data abstraction tool.

Variables. A patient flag in the IT system indicating need for laboratory testing; data from chart review on whether laboratory testing was ordered, was indicated, or was not indicated based on documentation; socio-demographic, and medical variables such as patient insurance, age, race, number of visits to CHC in last year, gender, other medical co-morbidities, and use of other medications.

Primary and Secondary Measures. We developed a machine query of the EHR at Baltimore Medical Systems (Maryland, USA) of patients eligible for the quality measure. Patients prescribed one or more of the target medications in the previous calendar year were included if they were 18 years or older, received at least a 180 day supply of the medication, and had not been hospitalized in the preceding year. Two hundred patients who were prescribed index medications (angiotensin converting enzyme inhibitors, digoxin, statins, or diuretics) were randomly selected. Eligibility for the quality measure (denominator) was determined by the query automatically and by two BMS employees manually probing the EHR. Reviewers extracted data on whether patient was seen in last 365 days, assess if documentation discusses

need for or orders the required therapeutic monitoring, and also assess potential exclusions to need for this therapeutic monitoring (e.g. patient was hospitalized in last 365 days, patient seen in external clinic or Emergency Department and had required therapeutic monitoring done there with laboratory results present in BMS chart, patient was discontinued from medication in question soon after 180 day criteria, etc). Discrepancies between manual reviewers and the electronic query were analyzed by a third reviewer who was knowledgeable of the dispute and adjudicated (gold standard).

Data Analysis and Outcomes. The primary outcome under Specific Aim 1 was to assess the accuracy of the EHR-based system in flagging patients needing therapeutic monitoring testing in comparison to the assumed bench mark standard of 80%. This information is crucial to determining if an approach of automatic reminders for overdue laboratory testing is even technically feasible with commercially available EHR systems. (In light of the new Meaningful Use requirements and quality reporting measures, the results in this section are extremely useful for determining the value of this quality measure in future stages of Meaningful Use.) With requirements of electronic submission of quality measures in Stage 1 of the Meaningful Use rulings, our experience with developing such an electronic query and reporting tool may provide some indication how challenging these requirements will be for private practices, who are not supported by the resources this grant provided us.

Based on the data collected, the accuracy of the EHR-based system was based on the positive predictive value (PPV) statistic, which was defined as the proportion of true positives (TP) as compared to the sum of the true positives and false positives (FP). By definition, a TP is counted if a patient is flagged for therapeutic monitoring by both EHR and chart review, while a FP is scored if a patient is flagged for testing by EHR, but not by chart review. In order to be relevant to clinicians, it was assumed that 80% PPV is necessary for EHR to be considered as a clinically viable alternative and we evaluated whether the PPV of EHR-based quality measures are significantly better or worse than 80%. Results are shown in the Results section.

Specific Aim #2

This aim includes the development and implementation of EHR-based quality and safety measures of medication monitoring for vulnerable populations that are served by BMS CHC that are useful to clinicians and senior leaders.

Data Sources/collection. To accomplish this Specific Aim, we developed and implemented via the EHR system Misys the four Medication Management quality measures. Patients flagged by this IT tool as in need of an annual therapeutic monitoring test for one or more of the quality measures had this 'quality lapse' fed back to their primary physician via a Performance Monitoring bulletin on approximately a bi-monthly frequency over the course of a year starting August 2008 (table 1), which is explained in detail under specific aim #3. With each bulletin, providers were surveyed about their intended actions to resolve overdue status for each overdue patient and their perceptions of the usefulness of the Bulletin as described above. Both the monitoring bulletins and provider surveys were distributed in the paper format at the 2 BMS sites, as web based survey mechanism was not found to be feasible.

In addition, opinions of three senior leaders (2 clinic medical directors and the Chief Medical Officer (CMO) of the system) on the bulletin and the project were collected via one anonymous survey and a face to face focus group.

Table 1: Distribution schedule for the monitoring bulletins and surveys (n = ~15)

Monitoring Bulletins and Surveys administration schedule	Distribution	Collection
Round 1	August 2008	September 2008
Round 2	November 2008	December 2009
Round 3	January 2009	February 2009
Round 4	March 2009	April 2009
Round 5*	April 2009	May 2009
Round 6	May 2009	June 2009
Round 7*	June 2009	July 2009
Round 8	July 2009	September 2009

* No surveys were distributed during these rounds to minimize respondent burden and the bulletins were shared to enable providers to take action they deemed appropriate.

Interventions. The surveys were developed to measure the complex concept of usability or usefulness. This target concept was selected as the least developed of the National Quality Forum's four criteria for evaluating quality measures: importance, scientific acceptability, usability, and feasibility. The concept of usability comprises ideas such as whether the intended audiences of a measure can understand the results of a measure, that differences in performance are clinically meaningful, whether audiences find the results useful for decision making, and whether the information produced by the measure actually results in a decision or action. Surveys sought to measure provider perceptions of usability or usefulness with a barrage of at least four items each assessing the contributing notions of timeliness, acceptability, understandability, and actionability. Responses ranged from 1-5 (Strongly agree to Strongly disagree) (a Likert scale). All items were reviewed in multiple rounds by an interdisciplinary group including seven physicians, two health services researchers, two nurses and several other staff of the community health centers.

Primary and Secondary Measures. Overall usefulness or usability; timeliness; acceptability (overall, of the bulletin itself, and of the medication monitoring quality measure); understandability; and actionability. Explanatory variables included gender, age, provider type, race, panel cap, number of overdue patients at the first time period.

Data Analysis and Outcomes. Our main a priori hypothesis with respect to the surveys was that 70% of frontline providers will describe reports of quality measure data as useful (defined as information that is delivered in a timely manner and be perceived to be acceptable, understandable, and actionable. In addition, we sought to understand what provider characteristics were associated with reports of usefulness and whether provider assessments of usefulness were related to provider action to be sure patients received medication monitoring.
Analyses included confirmatory and exploratory factor analysis to assess feasibility (preliminary convergent and discriminate validity) of scales. We plotted change in provider

attitudes over time by provider to assess how much variation in change compared to the group mean there was at the individual provider level (not shown). In addition, we conducted a nonparametric (small sample) tests of difference in group means at time 1 versus the end period (one year later) and several bivariate random effects regressions of usability and its factors on time (random effects accounted for repeated measures on providers). Finally, we conducted multivariate longitudinal analyses to examine associations of usefulness scores with 1) at the provider level, the percentage of a provider's patients overdue for medication monitoring who received monitoring by the next bulletin (within 2 months); and 2) at the patient level, whether the patient received monitoring by the next bulletin.

The second part of the a priori hypothesis that 70% of clinic directors and BMS senior staff would describe rolled-up versions of reports of quality measure data as useful, could not be tested as there are just 2 clinic managers and 1 overall director. In addition, all of these were also providers so their responses were included in the provider analyses.

Specific Aim #3

This aim involves the development and implementation of EHR-based quality and safety measures of medication monitoring for vulnerable populations that are served by BMS CHC that impact patient outcomes.

Data Sources/collection. To accomplish this Specific Aim, we developed and implemented via the EHR system Misys the 4 Medication Management quality measures. Patients flagged by this IT tool as in need of an annual therapeutic monitoring test for one or more of the quality measures had this 'quality lapse' fed back to their primary physician via a Performance Monitoring bulletin (paper format) on approximately a bi-monthly frequency over the course of a year starting August 2008 (table 1). 15 primary care providers received these monitoring bulletins at the 2 clinic sites. Providers received a bulletin every one to two months between August 2008 and August 2009 for a total of eight bulletins. Our original intention was to give providers a complete list of all of their overdue patients with each bulletin. We did so in August 2008 (bulletin round 1), but provider feedback suggested that the full list was too burdensome (number of patients listed ranged from 1 to 67 across all providers). As a result, for all subsequent bulletins (with the exception of bulletin round five (April 2009)), we showed only the 10 most overdue patients for each provider. In April 2009, we again provided complete lists of all overdue patients. Providers were not given specific instructions regarding how to manage patients on the bulletin but were given general information regarding the project's overall aim of improving medication monitoring.

In addition, the bulletin listed the patients most overdue for monitoring and their overall percentage of eligible patients, who had received annual monitoring, as compared to their peers in the same center. The intention for this performance comparison was to motivate "poorly-performing" providers to improve to peer-level.

Interventions. The medication monitoring bulletins was the main intervention. Every one to two months, a data analyst from the umbrella organization (HC) queried the EHR for eligible patients (automated query validated in specific aim#1 of the project). Patients so identified were aggregated based on their primary care provider (PCP) of record. Patients with no assignment to a PCP in the EHR were excluded. After consolidation of duplicate entries (e.g., one patient on

two index medications), patients were ranked based on the number of months overdue for recommended laboratory monitoring. Patients on more than one target medication were ranked based on the most overdue medication.

Following the data extraction, providers were sent a medication monitoring bulletin which included: 1) a summary of NCQA monitoring recommendations; 2) a list of the provider's overdue patients; and 3) a graphical summary of each provider's individual monitoring performance, each center's aggregate performance, and the overall performance of all providers at both centers. Providers reported their intended actions to resolve overdue status for each patient on their bulletin by choosing from a list of potential actions (e.g., "Sent patient a letter") and returned the bulletins to the research team. Providers could indicate patients that should be removed from future bulletins (e.g., "Patient no longer seen in the clinic").

Primary and Secondary Measures. The principal outcome measure was receipt of recommended laboratory monitoring. To assess the effectiveness of the provider-specific monitoring bulletin, a secondary outcome measure was whether or not patients identified as overdue for medication monitoring received recommended laboratory testing prior to delivery of the next bulletin i.e., patient level outcome of receiving monitoring. The exposure for this secondary outcome was whether or not the patient appeared on the monitoring bulletin at the time they were first identified as overdue.

Data Analysis and Outcomes. A single arm, pre/post intervention study was conducted using data from each of the rounds of bulletin distribution (table 1) at the 2 clinic sites. Demographic and patient characteristics for the 2,013 study eligible subjects were summarized and compared in bivariate tests by overdue status (overdue at least once during the study period versus never overdue) using the $\chi2$ test for categorical variables and ANOVA for continuous variables. To assess factors associated with subsequent receipt of recommended laboratory monitoring among patients identified as overdue, we restricted the logistic regression analyses to patients who had at least one episode of being overdue during the study period and non-missing data for overdue status at the time of the subsequent bulletin (N = 687). Each patient who either entered the study overdue for monitoring or became overdue for monitoring during the study contributed one observation to the data. For each patient, we analyzed whether or not that patient received testing prior to delivery of the bulletin in the next period after the patient was identified as overdue. For example, if a patient was identified as overdue at the time of the third bulletin, we assessed whether or not that patient remained overdue at the time of delivery of the fourth bulletin, with appearance (versus not) on the bulletin as the principal exposure. Models were adjusted for age, gender, clinical center, number of index medication classes prescribed, insurance status, primary care provider (physician versus mid-level), and number of months overdue at the time of the first overdue episode. The latter variable discriminated between those who were overdue at study onset and those who became overdue during the course of the study. All analyses were conducted in SAS, version 9.2 (SAS Institute Inc., Cary, NC).

Specific Aim #4

This aim evaluates the relationship between contextual factors (teamwork and safety climate at BMS CHC) and provider assessments of EHR quality and safety data as useful (Specific Aim

#2) and actionable (Specific Aim #3), as well as evaluating if the deployment of EHR quality and safety measurement efforts will improve teamwork and safety climate at CHC.

Data sources/Collection. To accomplish this Specific Aim we administered the Safety Attitudes Questionnaire (SAQ-Ambulatory Version) in three waves (table 2). The questionnaire was administered to both clinics, and all staff with a 50% or greater commitment to the clinic for at least the 4 consecutive weeks prior to survey administration.

Participation was voluntary and surveys were administered during pre-existing staff meetings, with a pencil and returned in a sealable envelope to maintain confidentiality. Individuals not captured in pre-existing meetings were hand delivered the same survey packet. This administration technique is known to garner an average response rate of over 80%, which is necessary for representative and interpretable teamwork and safety climate scale scores.

The administration of the survey at these 2 study clinic sites in this manner yielded expected results i.e., the administration got us an excellent response rate, which made interpreting the results much more straightforward (as was proposed originally in the grant). Namely the response rates for the 3 rounds of survey administration were:

Table 2: Schedule for administration of the SAQ

Safety Attitudes Questionnaire administration	Belair-Edison site	Highlandtown site
Round 1 (Quarter1, Year 1(October 2007))	83%	84%
Round 2 (Quarter 1, Year 2(September 2008))	84%	100%
Round 3 (Quarter 4, Year 2 (September 2009)	82%	95%

Each administration took 6 weeks to complete, with 3 weeks of preparation, and 3 weeks to administer. The administration timings of the SAQs provided 12 months of exposure to the intervention between the second and third administration the intervention could only be rolled out at the end of the 1st year due to IT interface related delays mentioned in the 'Challenges' section below).

Variables. Teamwork climate and safety climate are psychometrically validated scales from the Safety Attitudes Questionnaire, which elicit frontline worker assessments of collaboration and patient safety norms and practices, respectively.

Primary and Secondary Measures. SAQ scores of the 3 survey administrations at the 2 sites of CHC.

Data Analysis. Teamwork and safety climate scale scores at the respondent level were assessed over the 3 administrations of the SAQ to both the sites. To test for improvement in teamwork and safety climate assessments paired sample t-tests from baseline to post 21 months of exposure to the EHR data training, learning sessions, and use. This analysis was conducted using SPSS version 14.0 (Chicago, IL).A detailed review and discussion of the combined SAQ results of the three rounds with the clinical leadership and staff of both the sites was done in 2 separate sessions held after the project ended in November 2009. Results are mentioned in the Results section.

Challenges

Quality Measure Query related. Implementation of the national quality measure within the local EHR was very labor intensive and technologically challenging, but feasible. Despite highly experienced health information technology personnel, both at the FQHCs and on the research team, developing an automated system to produce data as specified by NCQA was extremely time intensive.(25) This raises questions about the feasibility of deploying measures such as these in unsupported ambulatory care practices with SIGNIFICANT implications for Meaningful Use quality reporting measures. Organizations that are developing quality measures should consider collaborating with larger EHR vendors to begin to develop "off-the-shelf" measurement tools, which might be easily deployed within existing EHRs in primary care for the purposes of robust quality measurement.

The providers in our study had minimal administrative time to conduct non-clinical tasks; no additional administrative time or support was given to providers during the study, potentially limiting their ability to effectively participate in this effort. Future iterations could potentially be strengthened by including patient- or pharmacy-focused interventions.

Limitations

The main, overall limitation of this study for analysis of provider-level data was the small sample size, since the study was powered on patients and not providers. Some more limitations linked specifically to specific aims of the project follow:

Safety Attitudes Questionnaire related. It was impossible to link the safety culture data to provider-level responses from other data collection. Further, we could not assess other events in the clinics that may have affected safety culture (such as through-put pressure, leadership changes, policy changes), perhaps in more important ways than our relatively small study.

Related to the EHR based Quality Measure. There are several potential threats to the validity of this study. First, this study evaluated EHR-based quality measure data to manual chart reviews. It is undoubted that manual chart reviews are not a 'gold standard'. This being true, manual chart reviews are currently the standard despite their imperfection and are the natural alternative to EHR-based quality measure data. We have taken this limitation into account in Specific Aim #1 and have been able to assess that not only EHR-based quality measure data is not inferior to chart review but also that EHR-based quality measure data is superior to chart review. By allowing the final arbiter to have access to the results of both manual as well as electronic review, the decision on accuracy could be done by this expert.

In an ideal world with limitless research funds, we would have liked to actually quantify the reliability of manual chart review for these quality measures. This approach was not feasible due to the enormous time and resources required and is, in fact, the main reason behind this grant, namely to evaluate whether EHRs can be reliably used for quality and safety measurement. Second, we realize our EHR efficacy for quality measurement is hampered by the fact that our demonstration site is a free-standing Community Health Center that, by definition, is not linked to a hospital to thereby strengthen the EHR data with aspects such as hospital visits or Emergency Department visits. We actually consider this to be strength however since demonstrating the utility of EHRs for quality data would be much easier in a clinic system linked

to a hospital database but the reality is that most ambulatory clinics are not readily attached to hospital databases, particularly ambulatory clinics that serve the most vulnerable populations like the CHCs. Partnering in the grant with such a clinic can predict the challenges that will be encountered in the implementation of Meaningful Use in many independent clinics.

   Patient data and Bulletin related. Since we were limited to data readily available in the EHR, some patient attributes of interest (e.g., race, income) were either absent or incomplete. While we examined the effectiveness of the provider bulletins in this paper, our study was principally designed to assess the feasibility of the quality measure implementation. A randomized, controlled design would provide more robust estimates of true bulletin effectiveness. In addition, while our data were generated in an automated fashion from the EHR, we were unable to integrate the paper-based provider monitoring bulletin into the EHR. The bulletin might have been more effective if it had been entirely EHR-based, facilitating direct documenting or lab ordering for overdue patients through links with the patients' electronic chart. Our non-randomized intervention focused, by design, on the most overdue patients. These patients may have differed in other, unmeasured ways from the less overdue patients not appearing on the bulletins. Similarly, the providers with the shortest lists of overdue patients were more likely to have newly identified overdue patients appear on their list (i.e., make the "top 10" cutoff). Therefore, apparent effectiveness of the bulletin might be partly attributable to characteristics of those providers most reliably performing medication monitoring at baseline, and not the bulletin itself. Finally, our effort was limited to two FQHCs. As such, our results may not be generalizable to other clinical settings. In addition, the inclusion of questions about intended actions and the survey of usefulness, required for the research study, may have increased burden for providers and reduced their time or willingness to use the bulletin to address patient needs.

Results

Specific Aim #1

   Principal Findings. The query dramatically outperformed manual reviewers compared to the gold standard (Table 3). Deficits of the query identified by the gold standard review resulted in query modifications subsequently resulting in 100% concordance with the gold standard.

Table 3.

	Electronic Query	Manual Review
Negative Predictive Value	99.4%	89.2%
Positive Predictive Value	96.9%	80.0%
Sensitivity	96.9%	37.5%
Specificity	99.4%	98.2%

   In the 7 cases where the gold standard did not match with the query, in 5 cases the inclusion/exclusion was not altered, and in 2 cases patients that had been excluded, had to be

included. The causes of discrepancy included missed insurance, missed new document type, and natural language processing problem related to ED visits.

Outcomes. The query dramatically outperformed manual reviewers compared to the gold standard (Table 3). Deficits of the query identified by the gold standard review resulted in query modifications subsequently resulting in 100% concordance with the gold standard.

Discussion. The efforts required to develop and validate this National Quality Forum Quality Measure in a electronic format casts a foreboding shadow on the quality measure requirements in the recent Meaningful Use rulings. While it is clear that electronic queries can outperform manual review (not only in accuracy but once developed with significant time savings), the development required in our case a team of system-analysts, informaticians, and IT managers. The total cost of developing such a query may be too high for most clinics.

Conclusions. Automatic queries of EHRs to identify patients eligible for quality measures are feasible and potentially far superior to manual reviews of EHR data. Complex eligibility rules for quality measures may limit the usefulness of human EHR reviewers.

Significance. Our findings hope to strengthen the evidence on the potential of EHRs for *Quality and Safety Measurement and Improvement* and thereby eliminate the enormous costs associated with using manual chart review to identify patients for quality measures such as annual therapeutic monitoring. The results of this grant will add to the several studies that have begun to explore the utility of Electronic Health Records (EHR) for quality measurement. One natural area of exploration has been the validity of use of EHRs to collect data on this annual therapeutic monitoring and whether this monitoring results in action by the provider and results in changes in patient outcomes. Existing studies highlight that despite terrific potential, EHR-based quality measurement is still in need of further research to understand its validity and ways to improve the validity. Our results highlight that the collection of data on quality measures via EHRs is possible and much less resource intensive if EHRs are already in place, and that the data collected is valid.

Implications. Automatic queries of EHRs to identify patients eligible for quality measures are feasible and potentially far superior to manual reviews of EHR data. Complex eligibility rules for quality measures may limit the usefulness of human EHR reviewers.

Specific Aim #2

Principal Findings. Factor analysis suggested the proposed scales required modification from the intended scales. The final overall scales and subscales had Cronbach's alphas ranging from 0.51 (timeliness) to 0.88 (overall).

For the survey of usefulness, response rates were more on the order of 73-80% for each reporting period. Averaged responses tended to be just above the neutral midpoint and were consistent over time. Random effects models of the usefulness scores over time found significant but small increases in provider assessments of the acceptability of the bulletin (as distinct from the quality indicator of annual monitoring) and the timeliness of the bulletin.

Our main *a priori* hypothesis with respect to the surveys was that 70% of frontline providers would describe reports of quality measure data as useful (defined as information that is delivered in a timely manner and be perceived to be acceptable, understandable, and actionable). However, this hypothesis was not supported. Only 1 to 5 providers ever scored the bulletin as useful (defined as the top two levels of a five-point Likert scale) on any subscale or the overall scale (representing between 8 and 50% of respondents in a given month.

Different provider characteristics were associated with the overall usefulness score and with usefulness subscales, controlling for the provider's score on the measure in the first time period. For example, providers who were African American found the bulletin significantly more useful than whites did in the overall and actionability scores (with a trend toward significance for acceptability of the bulletin). Providers of other race (compared with whites) found the bulletin less timely but more actionable (with a trend toward significance for understandability). Physicians found the bulletin significantly less actionable, less acceptable, and less understandable as compared to non-physician providers. Greater age was associated with a higher score for understandability.

Secondary hypotheses that the percentage of patients receiving monitoring after appearing on the bulletin might be associated with provider attitudes about the usefulness of the bulletin were not supported in provider- or patient-level multivariate analyses. This was true for the overall usefulness score as well as for all subscales. At the provider level, gender, provider type, age, and how many overdue patients the provider had in the first time period were all consistently associated with the percentage of patients per provider who received monitoring by the next bulletin. Female, older, and non-physician providers and those with a larger number of overdue patients in the first time period were more successful at getting patients monitored, controlling for attitudes about usefulness and for whether the full list of overdue patients was provided or just the top ten. Provider race and panel cap (top allowable size given provider type and full- or part-time status) were non-significant in virtually all models.

At the patient level, whether an overdue patient on the bulletin received monitoring by the next time period was related to other race, panel cap, provider type, and age, controlling for full list and provider attitudes about usefulness of the bulletin. Patients with providers of other race (as compared to white race) and physician providers were less likely to receive monitoring. Patients whose providers were older or who had large panel caps (e.g., with a full-time schedule) were more likely to receive monitoring.

Discussion. In general, the providers felt the surveys were burdensome and their reactions to having to complete bimonthly information may have dampened their enthusiasm for paying attention to the rest of the project. At one clinic, managers instructed providers to fully complete the intended actions on the bulletins and the surveys, within a very short turnaround time that was considered unreasonable.

Based on the clinic leadership's interest to make the project sustainable beyond its life, should it be feasible to continue, 2 BMS staff members who were also project team members were trained in case the sites wanted to continue the monitoring bulletin roll out after the project end date (August 2009) as an internal Quality Improvement effort. The 2 staff members were trained on the data extraction and bulletin preparation process during the last 4 months of the project. Despite doing this, it is not sure whether the bulletin distribution will continue due to provider time constraints.

Conclusions. The instrument used to measure providers' feelings about the usefulness of the feedback bulletin is potentially useful for measuring provider responses to other quality improvement or practice management interventions. However, the instrument needs to be tested further in larger samples. The provider response to the usefulness of the bulletin provided in this project was lukewarm at best.

Implications. Further efforts to offer feedback as a quality improvement tool should assess provider perceptions of usefulness in order to provide information on areas in which tools can be improved or to enhance provider engagement. Relative to the present project, however, other efforts to offer feedback to improve quality should reduce the data collection burden.

Specific Aim #3

Principal Findings. The 15 providers at the clinic sites (11 physicians, 4 mid-level providers) saw approximately 93,000 adult visits annually at the 2 centers combined. Providers responded to the bulletin with their intended action on a patient by patient basis about 95% of the time. The detailed results regarding their composition are given in a submitted manuscript.

Ever overdue versus never overdue. Across both centers, 2,013 patients were eligible for the study based on criteria previously described under the Methods section; of these, 1,164/2,013 (58%) were never overdue for recommended laboratory monitoring during the study period (table 4). Compared with patients who were never overdue for monitoring, those patients overdue at least once during the study were similar in terms of age, gender, and insurance. Patients on medications from 3 or more index medication classes were more likely to never be overdue for testing (67%), compared with patients on 2 (60%) or 1 (54%) medication class(es) (P < 0.001).

Table 4: Characteristics of patients prescribed index medications for at least 180 days

	Never overdue (N=1164), N (%)	Overdue at least once (N=849) N (%)	p-value	T overdue and T+1 resolved (N=110) N (%)	T overdue and T+1 overdue (N=593) N (%)	p-value
Female	799 (58.0)	578 (42.0)	0.69	81 (16.8)	401 (83.2)	0.22
Male	359 (57.1)	270 (42.9)		29 (13.2)	191 (86.8)	
Center = 1	579 (58.4)	412 (41.6)	0.59	42 (12.3)	299 (87.7)	0.018*
Center= 2	585 (57.2)	437 (42.8)		68 (18.8)	294 (81.2)	
Age (mean(std))	57.2(13.9)	56.3(15.4)	0.16	58.51(13.9)	56.03 (15.6)	0.12
No. Meds =1	558 (54.0)	475 (46.0)	0.0003**	54 (14.2)	326 (85.9)	0.52
No. Meds =2	453 (60.2)	299 (39.8)		45 (17.4)	214 (82.6)	
No. Meds >=3	153 (67.1)	75 (32.9)		11 (17.2)	53 (82.8)	
Overdue 08/08		294 (100)		57 (22.5)	196 (77.5)	0.003**
Not Overdue 08/08				37 (12.9)	251 (87.1)	
On tool for 1st overdue		392 (100)		66 (20.2)	261 (79.8)	0.002**
Not on tool for 1st overdue				44 (11.7)	332 (88.3)	

System-wide monitoring and the effect of the provider bulletin. At each round of bulletin, 74-79% of eligible patients in the two centers were up-to-date for medication monitoring, with a slight downward trend over time. Our study was not originally designed to assess the impact of the provider bulletins. Nonetheless, because the provider bulletins appeared to have no positive impact on the overall, system-wide performance on the quality measures, we conducted a post-hoc analysis of the medication monitoring of patients who were overdue during the study. Adjusting for other available variables, two were associated with receipt of testing prior to the subsequent bulletin: the center where care was delivered and appearance on the provider-specific bulletin. At Center B, where insurance coverage was lower and providers were less experienced, patients were less likely than Center A patients to receive testing (adjusted odds ratio (aOR) 0.6 [0.4-0.9]). Patients appearing on the bulletin were more likely than those not appearing on the bulletin to receive recommended testing (aOR 2.0 [1.3-3.1]). Whether a patient was overdue at the beginning of the study or became overdue during the study was not associated with subsequent receipt of testing (aOR for those more than 15 months overdue at identification (i.e., overdue at beginning of the study) compared with those 15 or fewer months overdue at identification (i.e., became overdue during the study) 0.8 [0.5-1.2]). Appearing on the bulletin at the beginning of the study had a positive effect on testing status.

Discussion. We found that two in five patients on ACEIs/ARBs, statins, digoxin and diuretics were overdue for laboratory monitoring during the yearlong study, based on NCQA monitoring recommendations. Individuals taking medications from multiple medication classes were less likely to be overdue. Provider-specific feedback reports increased the likelihood that identified patients would receive recommended testing. Limiting our intervention to only the most overdue patients, however, mitigated the overall impact of the intervention on system-wide monitoring performance. Further, other system events such as provider workload fluctuations, performance goals, policies had the potential to overpower the effect of our bulletin, which was one of many tasks providers had to complete.

Our finding that 58% of patients consistently received recommended monitoring during the study period is comparable to existing findings.

Implications. Prior research suggests that audit and feedback efforts can be effective in improving quality. In particular, feedback that is delivered with specific suggestions for improvement, frequently and in writing, may be most effective. However, moving beyond provider-focused interventions to multi-modal ones may further increase overall effectiveness.

Conclusions. Our study was not able to delineate the reasons why a subset of patients does not reliably receive timely medication monitoring, but there are likely to be many reasons, including patient-, provider- and system-related factors. Patient-related factors may include aversion to needle sticks, cost, and logistical barriers. Patients also may not perceive chronic medication use as risky or may assume they will become symptomatic if toxicities develop and plan to seek care at that time. Provider-related factors may include a lack of awareness that medication monitoring is indicated or a lack of belief in the value of medication monitoring. (33) System related factors may include technology inadequacies that make identification of patients due for screening difficult or impossible. Many clinical settings do not have fully integrated pharmacy, laboratory, and clinical data systems, which are a prerequisite for approaching medication monitoring in a robust, technology-driven fashion.

Significance. Our study found that individuals at risk for medication-related toxicity frequently go unmonitored. Provider-specific bulletins, based on nationally recommended quality measures, can enhance receipt of recommended monitoring among target patients, but implementation is technically complex and thus this type of intervention may not provide a robust fix at the population level. Optimizing the safety of patients using potentially hazardous medications will require further implementation research, likely involving multi-modal interventions.

Specific Aim #4

Principal Findings. The levels of staffing (as perceived by staff and providers) started very low in 2007 but trended toward improving at HC by 2009, while they started relatively normal in BE and trended toward worsening by 2009.

Both sites reported relatively low teamwork climate and safety climate in 2007, and these scores were not significantly different in 2009. In addition, Job satisfaction trended worse from 2007 to 2009 in BE, but did not change in HC.

Discussion. Although this was primarily an evaluation grant and not a Quality Improvement, based on the request of the CHC clinic site leadership, the results of the 3 SAQ survey administrations were fed back to the clinic leadership and teams in 6 debriefing sessions, 2 per year in each of the clinics within 3 months of analysis of the returned surveys (table 2). This enabled them to learn about their culture in real time, early on in the project rather than waiting until after the project was over. This was a positive deviation from what was proposed originally in the grant which and was done in the interest of the clinic taking initiatives to improve culture which in turn may impact patient outcomes (not measured as it was beyond the scope of this grant) early on rather than 2 years later.

Conclusions and Significance. Safety culture could be assessed in these clinics and the staff reported that the results did in fact represent their views of what it is like to work in the clinics. Taken together, these results indicate that the two sites where vulnerable populations were targeted also happened to be sites where innovation readiness was relatively low. This is reflected both in the safety culture results and in the face-to-face discussions with staff and providers during debriefing sessions. Staff and providers were struggling to get by against the backdrop of resource restrictions, increases in volume, and economic uncertainty. Caregiver resilience was low, creating an environment where HIT innovation, even if executed perfectly, was one more demand placed on the shoulders of staff and providers who were already struggling to do their old work, before the intervention was launched. The findings in safety culture are interesting in the light of lack of system-wide performance improvement in conjunction with the bulletin in aim 3.

To the extent that interventions designed to improve care quality are associated with improvements in safety culture from previous studies, we did not find evidence for that in the current study. This may be due in part to the limited impact that the intervention had on the staff and providers overall, and to possible selection bias. These clinics were assessed in the right timeframe, using the appropriate published methods, and generated excellent response rates. But both clinics had significant preexisting conditions prior to the implementation of the—namely,

their teamwork norms, safety norms, job satisfaction, trust in management, and staffing levels were relatively low. To put it in the words of a staff member "it is hard to be here, people are trying, but this is a tough time for lots of reasons.

    Implications. It is possible to examine the results by caregiver type within site, and this may be helpful for future research so that interventions in these settings, the caregivers involved, and the contextual factors such as safety culture, can be targeted in a consistent way. Perhaps a selection criterion for clinics to be used in future studies could incorporate innovation readiness, caregiver resilience, and protected time for project coordinators within the clinics.

List of Publications and Products

Papers

Bundy DG, Marsteller JA, Wu AW, et al. Medication monitoring in primary care: Putting national quality measures into practice. *Qual Saf Health Care.* Submitted August 2010.

Conference Proceedings

Lehmann CU, Bundy DG, et al. Automatic identification of patients eligible for quality measures using an EHR: feasibility and accuracy. 13th World Congress on Medical and Health Informatics; 2010 Sept 12-15; Cape Town, South Africa: MedInfo Congress 2010

Lehmann CU, et al. Impact of a Provider Notification Tool on Overdue Laboratory Testing. AMIA 2010 Annual Symposium, Nov 15, 2010; Washington D.C.

Lehmann CU, Bundy DG, et al. Automatic identification of patients eligible for quality measures using an EHR: feasibility and accuracy[abstract]. In: Safran C, Reti SR, Marin HF, Editors. *Studies in Health Technology and Informatics:* Proceedings of the13th World Congress on Medical and Health Informatics; 2010 Sept 12-15; Cape Town, South Africa: MedInfo 2010 IOSPress; Volume 160; 2010. P1367.