Grant Final Report

Grant ID: 1R18HS017048

Massachusetts Quality e-Measure Validation Study

Inclusive Dates: 09/12/07 – 08/31/11

Principal Investigator:
Eric C. Schneider, MD, MSc [*] [†] [‡] [§]

Team Members:
Carol Cosenza, PhD [**]
Jeffrey Linder, MD [†] [‡]
E. John Orav [†]
Tom Sequist, MD [†] [‡]
Steven R. Simon, MD, MPH [††]

[*] RAND Corporation, Boston MA
[†] Division of General Internal Medicine, Brigham and Women's Hospital, Boston, MA
[‡] Harvard Medical School, Boston, MA
[§] Department of Health Policy and Management, Harvard School of Public Health
[**] Center for Survey Research, University of Massachusetts, Boston
[††] Department of Ambulatory Care and Prevention, Harvard Pilgrim Health Care, Boston, MA

Performing Organization:
RAND Corporation

Project Officer:
Rebecca Roper

Submitted to:
The Agency for Healthcare Research and Quality (AHRQ)
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Abstract

Purpose: The purpose of this project was to evaluate the readiness of structured electronic health record (EHR) data in a community-wide, multi-payer health information exchange (HIE) to support ambulatory clinical performance measurement.

Scope: The Massachusetts e-Health Collaborative (MAeHC) implemented interoperable EHRs across hundreds of ambulatory practices in three diverse communities. The MAeHC-supported HIE created a Quality Data Warehouse (QDW) to gather key data elements on standardized quality measures and enable performance reporting to practices.

Methods: The project intent was to use data from the MAeHC Quality Data Warehouse to compare the availability of data elements necessary to support standardized quality measures to the availability using (1) a "hybrid method," combining claims data with medical record review in ambulatory practices, and (2) a "claims-only" method based upon claims data aggregated across commercial health plans and the Medicare program.

Results: A number of critical barriers prevented sufficient operation of the HIE over an adequate time frame to enable the comparative analyses envisioned in the project proposal. Even if HIE operations had been executed as intended, other critical barriers may have prevented successful evaluation of the adequacy of data for performance measurement in community-wide HIE systems. Evaluation barriers included the legal status of data sharing between HIEs and evaluators, conflicting Institutional Review Board (IRB) recommendations related to HIPAA regulation, timeliness of health plan legal review and data sharing, difficulty engaging patients in consent for evaluation of data, and challenges engaging ambulatory clinical practices in a difficult primary care environment.

Key Words: electronic health records, evaluation, health information exchange

Final Report

Purpose

The purpose of this project was to evaluate the readiness of structured electronic health record (EHR) data in a community-wide, multi-payer health information exchange (HIE) to support ambulatory clinical performance measurement. The project had two specific aims:

- Aim 1: For measures in the Ambulatory Quality Alliance (AQA) performance measurement set, to recruit a cohort of adult ambulatory patients from three communities that were piloting community-wide implementation of structured electronic health records (EHRs) to compare a quality measurement method based on a structured EHR data to a hybrid method involving a combination of aggregated claims data and medical record review.

- Aim 2: For measures in the AQA ambulatory care measurement set, using two secondary data sets on adult and pediatric patients in the same three communities to compare a measurement method based on structured EHR data to a "claims-only method" based on a novel database that aggregates claims data from commercial health plans and Medicare.

Scope

For the past two decades, performance measurement has assumed an increasingly prominent role in efforts to improve the quality of health care delivery. Despite the central role of performance measurement, the development of performance measures and reporting systems has been slow. In large part, this reflects the lack of easily retrievable clinical data in electronic format, which makes performance measurement a complex, cumbersome, and expensive task.[1] Moreover, the non-standardized nature of clinical documentation, billing, and claims data storage can lead to undesirable biases in comparisons of the performance of health plans, provider groups, and individual providers.[2] Despite these data-specific challenges, nationally standardized performance measurement sets have been endorsed by organizations like the National Quality Forum (NQF) and the Ambulatory Quality Alliance (AQA).

Electronic health records (EHRs) have the potential to improve the feasibility, reliability, and validity of performance measurement by improving the standardization and quality of the clinical data that are used to calculate measure results. However, adoption of EHRs has been slowed by financial, technological and other barriers. Even if EHRs are in wide use, their full potential for performance measurement might not be realized in the absence of integration of key data elements through health information exchange (HIE).[1] Comprehensive performance measurement of hospital and physician practice ideally makes use of exchangeable clinical data, derived from the disparate records held by ambulatory clinicians, hospitals, emergency departments, laboratories, pharmacies and health insurers.

The Massachusetts eHealth Collaborative (MAeHC) is a multi-stakeholder group formed with the aims of enhancing EHR adoption in Massachusetts, setting up clinical data exchange, and encouraging the adoption of computerized physician order entry (CPOE). Established with a coalition of 34 stakeholders who champion health information technology, and initial financial support from the Blue Cross Blue Shield of Massachusetts, the largest medical insurer in the state, MAeHC has been working to improve the quality and safety of ambulatory care in the state by supporting and demonstrating EHR adoption and use.

In 2005, the MAeHC completed a proposal process for selection of three Massachusetts communities as pilots for full statewide EHR implementation. The state of Massachusetts has over 6 million residents that are cared for by approximately 20,000 physicians in about 6,000 practices. The three selected communities represent over 150 ambulatory practices with more than 400 clinicians. As of 2009, the MAeHC had implemented full EHRs in all of these practices, and had created an operational clinical data exchange platform in two of the three communities.

A key opportunity of the MAeHC program was the development and testing of the community-wide health information exchange (HIE) to support performance measurement. While other large integrated delivery systems and medical groups have promoted use of EHRs within their systems and developed internal health information exchange to support performance measurement, the MAeHC initiative was among the few in the United States attempting to create community-wide, EHR-based, multi-payer HIEs that would enable aggregate performance measurement of independent ambulatory physician groups based on the full patient population of the group rather than simply assessing performance measurement of samples of patients with specific types of insurance. As proposed, the MAeHC initiative appeared to provide the ideal setting to assess the availability and quality of the clinical data elements in an HIE that would support ambulatory care performance measurement.

Methods

Overview of Design

The goal of Aim 1 was to compare the data available to measure ambulatory performance using the EHR and HIE approach with the data available using two traditional quality measurement approaches (the hybrid of claims data and medical record review and the claims data only). At the time the application was submitted for funding, the communities and EHR vendors had been selected and implementation of EHRs had begun, but the HIE, which depended on the EHR implementation had not yet occurred. We anticipated that the community HIEs would become operational on July 1, 2007. In addition, we selected a set of standardized ambulatory quality measures endorsed by the AQA and planned for implementation by the Massachusetts Health Quality Partners (MHQP) in collaboration with the MAeHC (Table 1).

Based on experience with a prior quality measurement project, we anticipated that privacy and confidentiality of medical data would require that patients consent explicitly to medical record review. Thus, we adopted a study design we had used previously to collect detailed clinical data for use in quality measurement. This involved a multi-stage data collection protocol. We would obtain a sample of patients from the community HIE, contact providers about the

survey, permitting them to "opt out" selected patients. We would survey the remaining patients to gather data about health care services that might not be well-represented in medical records. The mail survey was also an opportunity to invite patient permission to review patient health records and to carry out a consent procedure. In collaboration with the participating medical groups, we would arrange to conduct a medical record abstraction process for the patients who granted consent. Finally we would make requests to the HIE Quality Data Center (QDC) and health plans to obtain pertinent electronic quality measurement data and claims data respectively.

Table 1. Ambulatory quality alliance measures implemented via health information exchange in the Massachusetts e-Health Collaborative (MAeHC)

Table 1a. Asthma

Measure	Description	Age	Denom. Time Window	Numer. Time Window	Encounter Codes	Clinical Note	Diagnosis Codes	Med Codes	Lab Codes	Lab Results	Procedure/ Imaging Codes	Path/ Imaging Results
Use of Appropriate Medications for People w/ Asthma*	% of individuals, 5 to 56 years old, identified as having persistent asthma during the year prior to the measurement year and who were appropriately prescribed asthma medications (e.g. inhaled corticosteroids) during the measurement year.	18-56	2 yrs	1 yr	D		D	N				

Table 1b. Coronary artery disease

Measure	Description	Age	Denom. Time Window	Numer. Time Window	Encounter Codes	Clinical Note	Diagnosis Codes	Med Codes	Lab Codes	Lab Results	Procedure/ Imaging Codes	Path/ Imaging Results
Drug Therapy for Lowering LDL Cholesterol#	% of patients > 18 with CAD who were prescribed a lipid-lowering therapy (based on current ACC/AHA guidelines).	18+	1 yr	1 yr			D	N			D	D
LDL Cholesterol Level	% of patients with CAD whose most recent LDL cholesterol < 130 mg/dl or < 100 mg/dl (Using <100 for now pending PAG approval).	18+	1 yr	1 yr			D			N	D	D
Lipid Profile	% of patients with CAD receiving at least one lipid profile during the reporting year.	18+	1 yr	1 yr			D		N		D	D

Table 1c. Diabetes

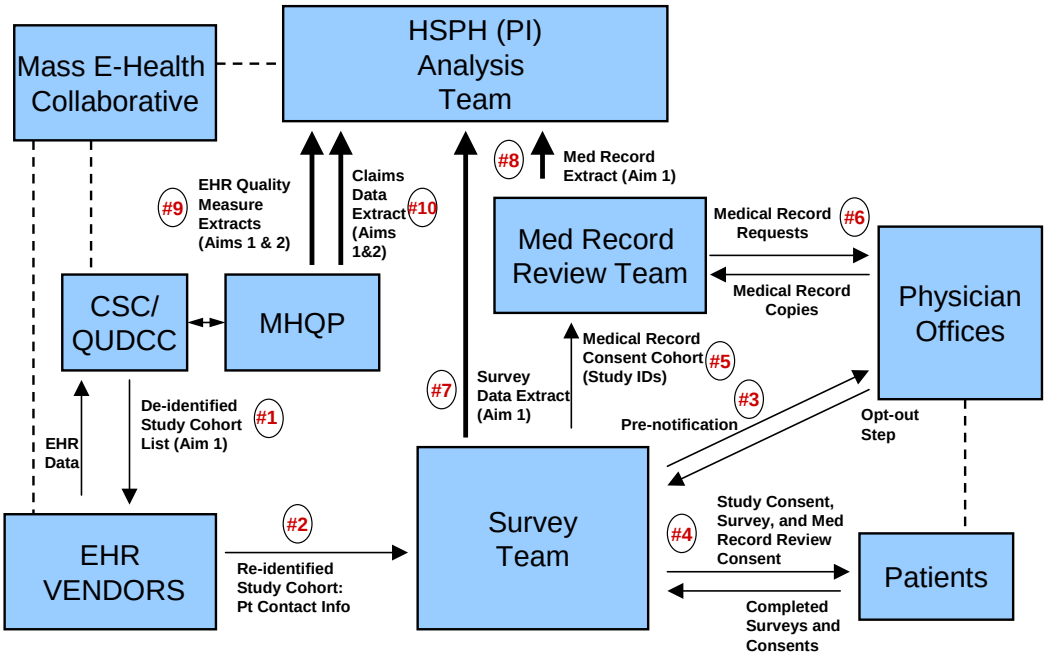
Measure	Description	Age	Denom. Time Window	Numer. Time Window	Enc-ounter Codes	Clin-ical Note	Diag-nosis Codes	Med Codes	Lab Codes	Lab Res-ults	Proc-edure/ Imag-ing Codes	Path/ Imag-ing Res-ults
Blood Pressure Management*	% of patients with diabetes who had their blood pressure documented in the past year less than 140/90 mm Hg.	18+	1 yr	1 yr			D	D				
Lipid Measurement*	% of patients with diabetes with at least one Low Density Lipoprotein cholesterol (LDL-C) test (or ALL component tests).	18+	1 yr	1 yr			D	D	N			
HbA1C Management Control*	% of patients with diabetes with most recent A1C level greater than 9.0% (poor control).	18+	1 yr	1 yr			D	D		N		
HbA1C Management*	% of patients, 18-75, with diabetes with one or more A1C test(s) conducted during the measurement year.	18-75	1 yr	1 yr			D	D	N			
Eye Exam	% of patients who received a retinal or dilated eye exam by an eye care professional (optometrist or ophthalmologist) during the reporting year or during the prior year if patient is at low risk for retinopathy. A patient is considered low risk if all three of the following criteria are met: (1) the patient is not taking insulin; (2) has an A1C less than 8.0%; and (3) has no evidence of retinopathy in the prior year.	18+	1 yr	1 yr	D		D	D		D	N	
LDL Cholesterol Level (<100mg/dL)	% of patients with diabetes with most recent LDL-C less than 100 mg/dL.	18+	1 yr	1 yr			D	D		N		

Table 1d. Prevention

Measure	Description	Age	Denom. Time Window	Numer. Time Window	Enc-ounter Codes	Clin-ical Note	Diag-nosis Codes	Med Codes	Lab Codes	Lab Res-ults	Proc-edure/ Imag-ing Codes	Path/ Imag-ing Res-ults
Influenza Vaccination	% of patients, > 50 at beginning of 1 year measurement period, who received an influenza vaccination.	51+	1 yr	1 yr							N	
Pneumonia Vaccination	% of patients > 65 who ever received a pneumococcal vaccine.	66+	1 yr	35 yrs							N	
Colorectal Cancer Screening*	The % of adults, 50-80 years old, who had an appropriate screening for colorectal cancer. One or more of the following: FOBT – during measurement year; Flexible sigmoidoscopy - during the measurement year or the four years prior to the measurement year; DCBE – during the measurement year or the four years prior; Colonoscopy – during the measurement or the nine years prior.	50-80	2 yrs	10 yrs			D				N	D
Breast Cancer Screening*	% of women, 42-69 years old, who had a mammogram during the measurement year or year prior to the measurement year.	42-69	1 yr	2 yrs							N	

The overall design involved separate research teams collecting specific data (claims, survey, medical record data, EHR data) with the lead team creating a study identifier unique to each patient to enable patient-level linkage of the separate data sources. The intent of this was to minimize the risk of inadvertent disclosure of patient identity with other forms of personal health information (Figure 1). Analytic files combining the data were to be constructed using the unique study identifier.

Figure 1. Overview of data collection protocol



Instruments

Measure selection: Of the 26 AQA-endorsed quality measures, 14 measures pertaining to adults were specified by the MAeHC team for implementation in the Quality Data Center (QDC) based on EHR data feeds through the HIE. We developed a comprehensive, detailed specification of the data elements and all of the plausible data sources (EHR, paper medical record, patient report, and claims) corresponding to these measures (Table 1). Based on this comprehensive data map, we developed an adult patient survey capable of capturing relevant data elements and a medical record abstraction tool that could be used to gather relevant data elements from on-site review of medical records (including both paper-based and electronic health record sources). The data collection tools along with the invitation and consent documents were reviewed and approved by multiple IRBs.

In the original proposal, the timing of the MQeVS data collection was dependent on three key precursor activities: (1) development of the capacity to sample patients from the Community Health Information Exchanges (HIE) that supply clinical data to the Quality and Usage Data Coordinating Center (QUDCC) electronic data repository (which was ultimately renamed the "Quality Data Center" or QDC); (2) accrual of patient data in the QDC; and (3) implementation and testing of electronic specifications for the AQA quality measures using the QDC.

Sample

Two health plans (the two largest health plans in Massachusetts) agreed to provide rosters of patients with contact information in the study communities. Furthermore, physician groups in the largest of the three communities were not able to come to agreement about the electronic

health record vendor and they were not able to establish an HIE. To address this challenge, we approached the two largest health plans and asked them to supply rosters of potentially eligible patients living within the two MAeHC communities with HIE.

We requested that health plans use the claims data to flag patients with specific chronic conditions (asthma, diabetes, heart disease) so that we could oversample patients likely to be included in the denominators of relevant chronic condition quality measures. Citing HIPAA regulation, health plans were not willing to release diagnosis information without patient consent. Lacking these data, we were limited to sampling patients at random from these rosters.

To enrich the sample with patients who would be likely to have accrued data in the HIE, we requested and received a roster of MAeHC participating physicians with summaries for each physician of the number of "messages" shared with the HIE in each community. We sampled patients of the providers with the largest number of messages entered into the HIE. The two health plans provided rosters of 2,949 patients who made 4,041 visits to MAeHC-participating primary care providers (designated based on specialty). In order to even the burden of medical record review, we capped the total sample from each provider at no more than 175, selecting patients at random for those providers with >175 patients. If a provider had fewer than 175 patients, we took a census of all of those patients. Although our goal was 2,100 patients, after restricting the sample to patients living within 50 miles of a provider and applying the restriction based on a minimum number of messages shared with the MAeHC HIE, the final eligible sample included 1,919 patients (who had at least one visit to one of 17 providers).

Data Collection

Based on the initial eligible sample of 1,919 patients, we began the survey fieldwork and consent procedure. The details are presented in the results section below. For patients who provided consent to medical record review, once we had received all of the consents, we contacted providers and arranged for on-site medical record abstraction. While this medical record abstraction process was under way, the second no-cost extension period ended. We requested a third no-cost extension, but this request was not granted. We halted data collection under this grant.

For Aim 2, the failure of the HIE to be implemented in the largest of the three MAeHC communities, the small number of providers with adequate message submission to the HIE, and the requirement that providers request the "re-identification" of patients in order for MAeHC to share QDC data meant that our planned approach to comparing the health plan claims data with the data obtained from the QDC was not going to be linked in a way that would yield meaningful comparisons of the data element availability from these two data sources. Specifically, measures applicable only to patients with chronic diseases would involve very few individuals rather than the thousands that we had anticipated had the HIE been fully operational in the three communities. The alternative plan for Aim 2 (obtaining the claims and QDC data for the patients who provided written consent as part of the survey) was not achievable because the project was terminated before we could engage providers in the re-identification procedure.

Results

This section summarizes key challenges that in the end prevented the research team from carrying out the proposed evaluation. We are unable to report the anticipated quantitative results related to the two aims of the study. Nevertheless, the project has produced numerous insights into the strategy for evaluating the availability and quality of performance data from community-wide HIE including tools for such an evaluation and a lengthy list of challenges that will confront future evaluators seeking to carry out similar evaluation work.

Inadequate implementation of the HIE function plus additional challenges noted above left the research team with limited numbers of eligible patients, practices, and doctors to provide the data necessary to accomplish the study. EHRs were adopted in the MAeHC communities, but our project on performance measures required implementation of functional health information exchange (HIE) between EHRs. In the original proposal, we foresaw that implementation of HIE might be incomplete. To protect against partial implementation of HIE, we employed a sampling strategy that would enable compensating adjustments. Furthermore, we proposed sampling and data collection methods that had proven successful in previous projects. Nevertheless, new vulnerabilities of these methods related to IRB review and varying interpretations of HIPAA regulation slowed the progress of data collection. The extent of the HIE implementation was not apparent until we began the survey fieldwork. Taken together these challenges prevented us from producing meaningful quantitative analysis of performance data generated by the MAeHC.

We summarize two sets of results. The first set of results concerns the evaluation framework, operational plan, and tools that can be used by future evaluators of performance measurement in the context of community-wide (or institutional)HIE. The second set of results summarizes important barriers to future evaluators who may seek to test the validity and reliability of performance data derived from electronic health records (EHRs) using HIE. These latter results can inform future studies involving comparison of HIE data with other types of data used in current performance measurement programs (e.g., claims data, medical record abstraction, and patient survey).

Evaluation Framework, Operational Plan, and Tools

Adequate assessment of the validity of performance data requires meaningful comparator data to evaluate the specifications of a measure as implemented using different data sources. Traditional comparisons of EHR-based performance measurement evaluate measure results calculated using EHR data with measures results calculated using claims data, paper medical records, survey data or other data sources.

Our proposed evaluation framework anticipated that standard data elements could serve as a common set of “building blocks” to be used to specify multiple performance measures. Instead of focusing on a measure-by-measure analysis, we developed a framework focused on the evaluation of the data elements likely to be common to multiple measures. For example, the diagnosis of “diabetes” may be used in multiple performance measures based on varying data sources. The data elements used to define a patient with diabetes will differ across EHR, claims, paper medical records, and survey. The evaluation framework recognizes this data source variability and addresses the possibility that HIE may have to incorporate varying specifications

of diabetes captured by different EHRs (either because EHRs vary in the collection and storage of this concept or clinicians vary in their use of a standard code). The evaluation framework seeks to assess the magnitude of the bias introduced by specification and data differences at each level by creating a common specification of the construct "diabetes" across different data sources.

A second issue for evaluation of HIE for performance measurement is the effect of a transition of clinical office practices from a wholly paper-based medical record keeping system to an increasingly electronic record keeping system. Office-based practices that are in a "hybrid" state, using both EHR data and paper records might vary considerably in their manner of storing and transmitting data to the HIE. Some office medical records may be fully electronic while others may continue to store some current (and historical) clinical data on paper.

The third key issue for the evaluation of HIE is the potential for loss of information at the time of transmission from the EHR to the HIE. It is commonly believed that data in an HIE will accurately reflect the data transmitted from EHRs. However, it is not clear that this will be so when EHRs from different vendors are transmitting data into an HIE. Thus, comparison of the data represented in an HIE with the data represented in an office EHR is necessary to test the hypothesis that the HIE data maintain fidelity with the EHR data.

To address these three issues, we created two types of data collection and analysis tools. The first set of tools is the MQeVS Data Element Cross-walk. An example of this cross-walk for a single measure (control of blood pressure) is shown in Table 2. It summarizes each of the data elements required to specify measures and details how each data element is specified in a cross-walk of each of four data sources: Administrative data (or claims), medical record (as abstracted manually from either EHR or paper medical record), patient survey, and the HIE (known as the QDC in Massachusetts). The Cross-walk defines a variety of "types" of data elements that would be required for implementation of any measure. This assures consistency of specification among measures (with building blocks defined by the rows of the tables). It also specifies the basis for each data element across the data sources enabling comparison of the specifications (across the columns of the tables).

Table 2 shows the Data Element Cross-walk that we specified for only one of the measures included in the project. A similar cross-walk including other types of data elements was generated for each of the measures included in the project. This cross-walk enabled us to design patient survey items, medical record abstraction protocols, and administrative data queries that maximized consistency of specification across data sources. While the cross-walk presented in Table 2 is specific to an ambulatory performance measure, this cross-walk is designed to be applicable to any quality measure and potentially to any number of data sources that might be used in a comparative analysis.

Table 2. Example: Data element cross-walk for measure of blood pressure control

Condition: Hypertension
Measure: Blood Pressure Controlled
Description: The percentage of adults, 18-85 years old, who had a diagnosis of hypertension and whose blood pressure was adequately controlled (<140/90) during the measurement year
Note: The numerator and denominator time windows listed below may be adjusted at the time of analysis.

Measure: BP Controlled	**Administrative** *****each code has an associated date **Developer:** NCQA **Version/Year:** 2008 **Source Document:** HEDIS 2008 Tech Specs Vol 2	**Patient Survey** (Version 9/3/2009)	**QDC (Quality Data Center** **Source:** Quality Measures Specifications, August 2008	**Medical Record Abstraction** (version 9 9-24-2009) **Bold =** Items in HEDIS Hybrid (also noted) **Underlined =** Items in HEDIS Admin only **Regular =** Non-HEDIS Items for Bronze Standard/Additional Analyses
Age	DoB	#L1 (year of birth)		**#14 (DOB)** **#45 (inclusion age)** (in HEDIS Hybrid)
Sex	m/f	#C22 (m/f)		#13 (m/f)
Den Time Window	1yr	1yr	1yr	*** NOT SPECIFIED ***
Num Time Window	1yr	1yr	1yr	
Encounter - Denominator Inclusion	**Outpatient:** (CPT) 99201-99205, 99211-99215, 99241-99245, 99384-99387, 99394-99397		**At least 2 face-to-face visits with clinician** (CPT) 99201-99205, 99212-99215, 99241-99245, **99341-99350, 99354-99355, 99385-99387, 99395-99397, 99401-99404, 99411-99412, 99420-99429**	
Encounter - Numerator Inclusion				#49 (Visit with provider for HTN)
Patient Symptoms or Characteristics - Denominator Inclusion			Age 18-85 as of the 1st day of the measurement time period	#45 (inclusion age)
Diagnosis Denominator Inclusion	**Hypertension:** (ICD-9-CM Diagnosis) 401		**Diagnosis of Hypertension** any time in patient's history (ICD-9-CM Diagnosis) **410.0,** 401.1, 401.9, **402.xx, 403.xx, 404.xx**	**#46 (Dx hypertension)** (in HEDIS Hybrid)

Table 2. Example: Data element cross-walk for measure of blood pressure control (continued)

Measure: BP Controlled	**Administrative ***each code has an associated date **Developer:** NCQA **Version/Year:** 2008 **Source Document:** HEDIS 2008 Tech Specs Vol 2	**Patient Survey** (Version 9/3/2009)	**QDC (Quality Data Center Source:** Quality Measures Specifications, August 2008	**Medical Record Abstraction** (version 9 9-24-2009) **Bold =** Items in HEDIS Hybrid (also noted) **<u>Underlined =</u>** <u>Items in HEDIS Admin only</u> **Regular =** Non-HEDIS Items for Bronze Standard/Additional Analyses
Diagnosis Denominator Exclusion	**Evidence of ESRD:**		**End-Stage Renal Disease:**	
	(CPT) 36145. 36800, 36810, 36815, 36818, 36819, 36820, 36821, 36831-36833, 50300, 50320, 50340, 50360, 50365, 50370, 50380, 90920, 90921, 90924, 90925, 90935, 90937, 90939, 90940, 90945, 90947, 90989, 90993, 90997, 90999, 99512 (HCPCS) G0257, G0308-G0313, G0314-G0319, G0322, G0323, G0326, G0327, G0392, G0393, S9339 (ICD-9-CM Diagnosis) 585.5, 585.6, V42.0, V45.1, V56 (ICD-9-CM Procedure) 38.95, 39.27, 39.42, 39.43, 39.53, 39.93-39.95, 54.98, 55.6 (UB Revenue) 0367, 080x, 082x-085x, 088x (DRG) 317 **Pregnancy:** (ICD-9-CM Diagnosis) 630-677, V22, V23, V28		(CPT) 36145, 36800, 36810, 36815, 36818, 36819, 35820, 36821, 36831-36833, 50300, 50320, 50340, 50360, 50365, 50370, 50380, 90920, 90921, 90924, 90925, 90935, 90937, 90939, 90940, 90945, 90947, 90989, 90993, 90997, 90999, 99512 (HCPCS) G0357, G0314-G0319, G0322, G0323, G0326, 0327, S9339, G0308-G0313 (ICD-9 Diagnosis) 585.5, 585.6, V42.0, V45.1, V56 (ICD-9 Procedure) 38.95, 39.27, 39.42, 39.43, 39.53, 39.93-39.95, 54.98, 55.6 (UB-92 Revenue) 0367, 080x, 082x-085x, 088x (DRG) 317 **Pregnancy:** (ICD-9 Diagnosis) 630-677, V22, V23, V28	**#47 (Evidence of ESRD)** **#48 (Evidence of pregnancy)**
Pharmacy Inclusion		#E2 (currently taking meds to control your BP?)		**#50 (meds for high BP)**

Table 2. Example: Data element cross-walk for measure of blood pressure control (continued)

Measure: BP Controlled	**Administrative** ***each code has an associated date **Developer:** NCQA **Version/Year:** 2008 **Source Document:** HEDIS 2008 Tech Specs Vol 2	**Patient Survey** (Version 9/3/2009)	**QDC (Quality Data Center) Source:** Quality Measures Specifications, August 2008	**Medical Record Abstraction** (version 9 9-24-2009) **Bold =** Items in HEDIS Hybrid (also noted) **<u>Underlined =</u>** <u>Items in HEDIS Admin only</u> **Regular =** Non-HEDIS Items for Bronze Standard/Additional Analyses
Procedure/ Imaging Occurences - Numerator Inclusion		#E1 (ever been told by a doc that you have high BP?)	Obtain most recent blood pressure reading from within measurement year on or after when diagnosis of hypertension was made. Failure to document will be considered failure to treat	
		#E3 (in last 12 months, did you have your BP checked?)	**Recording Blood Pressure:**	**#51 (BP reading)**
			Systolic (SNOMED) 271649006 (CPT Cat II) 3076F, 3077F, 2000F*	
			Diastolic (SNOMED) 271650006 (CPT Cat II) 3078F, 3079F, 2000F*	
		#E5 (was your BP high, borderline, normal or low?)	**Blood Pressure Level:** (both systolic and diastolic are required)	**#51a (BP results, dates, sources)**
			Systolic (CPT Cat II) 3074F, 3075F	
			Diastolic (CPT Cat II) 3078F, 3079F	
Procedure/ Imaging Occurences - Numerator Exclusion			**Blood Pressure Level:** (both systolic and diastolic are required)	**#51 (BP reading)**
			Systolic (CPT Cat II) 3077F	**#51a (BP results, dates, sources)**
			Diastolic (CPT Cat II) 3080F	
Other				#7 (Pt ID) #8 (managed care ID) #9 (abstractor code) #10 (Abstraction date) #11 (Physician Name) #12 (Physician Practice) #15 (ethnicity) #16 (Race) #17 (Marital status) #18 (zip) #19 (Date of last outpatient visit note) #20 (Date of last history and physical note)

The second set of tools encompassed the data collection instruments specific to each data source. Based on the data element definitions included in the measure-specific data element cross-walks, the data collection instruments included (1) the administrative data codes used to define each concept in claims and administrative data (based on AQA, HEDIS, or PCPI measure specifications as appropriate), (2) the medical record abstraction tool including the definition of data elements to be gleaned from medical records, and (3) the patient survey including the survey questions that would be asked of patients to determine, for example, whether a diagnosis was present or whether a service was provided.

Two data collection instruments (the medical record abstraction instrument and the patient survey instrument) were designed and implemented specifically for this project. The other data collection instruments were algorithms specified by other organizations that were to be refined using the data we collected. For example, the third data collection tool was the algorithm specifying the data elements for each measure as created by the MAeHC QDC. The fourth data collection tool was the administrative measure specification developed by the National Committee for Quality Assurance (NCQA) or the American Medical Association Physician Consortium for Performance Improvement (AMA PCPI). As noted, we did not refine these specifications because the project was terminated before the algorithms could be tested using QDC and health plan administrative data.

Barriers to Evaluating Performance Measurement in the Context of Community Health Information Exchange (HIE)

The evaluation of an HIE is critically dependent on the set of agreements that govern its operation. These agreements include multiple stakeholders in the local community, some of whom are competitors. The agreements define the success of both development and operation of the HIE. They also constrain what may be evaluated. In particular, agreements with patients about the conditions for use of their data by the HIE can be especially problematic from an evaluation standpoint. This section describes the multiple manifestations of these fundamental challenges to evaluation of community-wide HIE and discusses some mitigation strategies that might be used in the future.

1. Slow and Incomplete Implementation of Community Health Information Exchange (HIE). The implementation of the community HIE by the Massachusetts eHealth Collaborative is probably typical of the experience of most community-wide HIE implementations involving multiple independent organizations. The fundamental challenge is the requirement to establish agreement among multiple organizations (hospitals, physician groups, and others) who will both feed and retrieve data to and from the HIE. These stakeholders are often competitors in a changing health care marketplace. The incentives for HIE are not necessarily strong and there is limited urgency. In Massachusetts, even in communities that created HIEs, the establishment of the HIE was delayed by nearly 18 months. The largest of the three community HIEs was never implemented because of disagreements among local stakeholders.

2. Technical Problems with HIE. Because HIE data standards are relatively primitive, technical problems in data sharing are likely, especially during early phases of implementation and if multiple vendor EHRs are included. Shortly after it began to operate, one community HIE had to be shut down because of technical issues matching laboratory results with laboratory

orders correctly. Such problems can occur in any data management environment, but may be especially prevalent when exchange of data among multiple organizations using different vendor systems can increase the risk of data transmission errors. The implication for our project was that none of the community HIEs was able to accrue clinical data over a sufficiently long time to permit adequate performance measurement. Most performance measures require at least one year of data and many require at least one additional year of "look-back" to establish inclusion or exclusion of a patient based on diagnoses or procedures.

3. Legal and Regulatory Barriers to Use of HIE Data for Evaluation. To evaluate the validity and reliability of HIE data for performance measurement, data from alternative sources such as patient survey, medical records, and health plan claims should be combined. However, under HIPAA retrieving these other data types with the necessary personal health information (PHI) requires study participants to give consent. Legal issues related to data sharing and health information exchange are especially challenging in a community-wide HIE because multiple organizations have jurisdiction over some part of the data and process (as opposed to research in single types of organizations such hospitals, medical groups, or health plans). Substantial resources are required to obtain clearance for data sharing to support evaluation. These resources cover applications to all of the relevant organizations including not just multiple Institutional Review Boards (IRBs) with oversight, but also HIE legal counsel, health plan legal departments, and HIE vendor legal counsel.

4. Lack of Consistency across IRBS and Other Legal Entities Reviewing and Overseeing the Data Evaluation Protocol. Our project encountered a heightened sensitivity of Institutional Review Boards (IRBs), MAeHC legal counsel, health plan legal departments, and potential study subjects to the requests to collect personal identifying data that would enable recruitment and clinical data that would enable the evaluation. This heightened sensitivity may have been related to a small number of high-profile breaches of electronic data files reported in the national news media during the time we sought to resolve these issues, but in the absence of standards, each entity identified additional barriers and requirements that ultimately had to be submitted to others that had already approved the protocol.

The sheer number of reviews of recruitment materials, surveys, and consent forms for this highly complex data collection project generated dozens of requests for clarification and modification, particularly from multiple IRBs and health plan legal departments. Even minor modifications to forms required re-review by the three IRBs involved in project oversight. Ultimately, the multi-stage process for obtaining consent, approval, and establishing data use agreements with many organizations was simply too cumbersome to be practical. As an example, late in the project, after the survey and consent fieldwork was completed, a physician group questioned whether the consent form approved by three IRBs and 2 health plans was in fact HIPAA compliant. On re-review, the three IRBs admitted that they had erred in approving the original form. The research team was left with the choice of obtaining a post-hoc waiver from the IRBs or going back to the research subjects to obtain signatures on revised consent forms.

Among the potential solutions to legal and regulatory challenges would be to include in the patient's original consent permitting use of the EHR data for HIE the language necessary to permit evaluators to approach patients for the purpose of evaluation of the accuracy of data included in an HIE.

5. Challenge to Engaging Patients in Evaluation of HIE. An important barrier in the HIE setting is that the consent given by patients for sharing of data with the MAeHC-sponsored HIEs did not permit the sharing of the names and contact information of individual patients with anyone other than treating physicians and then only by special request. Furthermore, the protocol that physicians had to follow to "re-identify" patients involved submitting a single request for each patient. This formidable barrier made the recruitment of patients to the study extraordinarily challenging and we ultimately opted to draw the sample from health plans because they had established mechanisms for allowing researchers to invite patients to participate in evaluation.

6. Sampling Challenges. One advantage of a community-wide HIE for performance measurement is the capability to assess the quality of care for samples of patients regardless of the type of health insurance or the providers that they visit. For reasons noted earlier, the project recruited community-based cohorts of patients through two large health plans. Nevertheless, we faced a severe constraint on the available sample due to the challenges that are listed below. The cumulative effect of these challenges was a sample far below our original projections. According to original projections obtained from the Massachusetts eHealth Collaborative (MAeHC) at the time of the proposal, 425 physicians (of whom 178 physicians were primary care physicians) were participants in one of the three MAeHC communities. In the proposal, we projected that if only 10% of the patients of these physicians were eligible for the study, this would provide 22,000 eligible patients.

- Sample Challenge 1: Incomplete participation in HIE. Physicians in the largest MAeHC community, Brockton, were unable to implement the community HIE and therefor were necessarily excluded from the sample. In addition, one of the three health plans dropped out of the study and could not be persuaded to rejoin. However, this was the smallest of the three health plans and the effect on sampling from our three communities was expected to be modest.

- Sample Challenge 2: Reliance on two health plans to provide a commercial insurance sample. The approach excluded patients with Medicare, Medicaid, and the uninsured as well as the hundreds of other commercial insurance plans. Furthermore, it proved difficult for insurers to identify only those patients who lived within the MAeHC communities and to exclude patients who were living primarily in other states, but had contact with providers in the MAeHC communities.

- Sample Challenge 3: Inability to selectively sample patients with chronic disease prior to recruitment. To enhance the efficiency of the analysis, we planned to identify members eligible for one or more HEDIS measures by virtue of their eligibility for preventive services or having a chronic condition that is the subject of a quality measure (diabetes, CAD, asthma). This would permit us to oversample those members who are most likely to have data in the HIE while excluding those who would not. Citing HIPAA concerns, the participating health plans refused to provide samples of the members known to have specific chronic health conditions along with the member contact information.

- Sample Challenge 4: Low physician submission of data to the HIE. A major unanticipated challenge to our sample was in the low participation of physicians in the

MAeHC-sponsored community HIE. Although the number of physicians using EHRs in the two remaining study communities was adequate, HIE message counts were low. These message counts provided to us by MAeHC staff in late 2010 indicated the degree to which data were being entered into the HIE. Based on our analysis only 17 physicians had a sufficient number of messages transmitted to the HIE to populate one or more of the quality measures of interest. Nevertheless, there were more than 3,000 eligible patients in the two health plan rosters among those physicians. We sampled 2100 patients as planned.

7. Incomplete Cooperation of Community Physicians. Among the 17 physicians approached for inclusion in the study, one large group of 7 physicians opted out all of the health plan members on their list before we could invite the members to participate. We followed up, but were unable to persuade the leader of this primary care group to participate despite offering to cover all expenses, provide needed staffing, and pay an honorarium to the group to encourage their participation. One physician had recently left the group and the group leader did not feel they could cover existing demand from patients and coordinate participation in the study. The loss of these 7 physicians cut our overall sample to 1,313 eligible health plan members despite replacement of sample from other physician practices.

8. Low Response Rates to Health Plan Member Survey. Based on earlier experience with the same protocol, we anticipated a 60% response rate with an 85% rate among respondents of returning consent forms for medical record abstraction. Despite adding time and follow up prompts, the final achieved survey response rate was 34.5% (N=446) and the final rate of completion of consent forms was 21.4% (N=276). While troubleshooting the low response rate during fieldwork, we identified a number of potential contributing factors. First, the health plan was identified as the source of the data. A small number of inquiries by potential respondents suggested that not all health plan members were comfortable having their contact information shared with an external evaluation team. The contribution of this problem to the low response rate is impossible to quantify. Second, the consent form and explanatory materials required by the health plan legal departments and the IRB were extensive, complex, and highlighted risks using technical language that respondents may have found troubling even though the chances that such problems would occur were remote. Third, survey response rates have been dropping in general since the previous use of this protocol. A 60% response rate may have been optimistic. Fourth, a number of high-profile medical record data breaches occurred just before and during our fieldwork, receiving national attention in the media. Some respondents expressed concern about these news stories leading us to believe some patients were nervous about providing consent for researchers to review medical records. We observed a higher than expected rate of active opt-outs and refusals to both the survey and the consent form compared to prior studies of this type. Fifth, the migration to cell phones and the use of call-screening have reduced the yield of telephone follow up that has been used in the past to enhance survey response rates.

Consequences of Logistical Challenges for the Project

The exceedingly low penetration of the community HIE and the subsequently high attrition rate among patients recruited to consent for medical record review affected both aims of the project. Without the ability to oversample patients with chronic disease, the samples of patients reporting chronic conditions (the population sampled for chronic disease management performance measures) were too small to produce statistically meaningful results for any of the data elements of interest. Table 3 summarizes the final survey response sample including the preliminary numbers of individuals eligible for the denominators of performance measures for specific conditions.

Table 3. Maximum denominator sizes for clinically-defined subsets (Survey N=446)

Condition/Measure (age range)	Number of Respondents	Percentage
All (18+)	446	100.0
Influenza vaccination (>50)	265	59.4
Pneumonia vaccination (>65)	32	7.2
Colorectal cancer screening (50-80)	279	62.6
Tobacco use (18+)	439	98.4
Breast cancer screening (women 42-69)	174	39.0
Hypertension	156	35.0
Coronary Artery Disease	35	7.8
Diabetes	40	9.0
Asthma	37	8.3

While we established the feasibility of the protocol in part, the final sample of 276 patients completing the survey and providing consent for medical record review was less than one quarter of the anticipated sample. The final analytic sample required medical record review and this proved challenging logistically. For 109 records, physician practices declined to participate. Among 123 patients whose clinicians agreed to participate, only 72 records could be retrieved and abstracted. For the other patients, offices did not have records of qualifying visits during the study period. The remaining 44 records of patients who consented could not be retrieved because the project was terminated before abstraction could be scheduled. Because the disease-specific and sex-specific quality measures are subsets of that sample, the power of the study to detect differences in data availability between modes of data collection would have been extremely limited (as Table 3 demonstrates).

Aim 2 was based on a comparison of the quality of care of practices in the communities measured by the MAeHC QDC and by health plan claims data. Because of the very limited set of EHR quality measure data on 17 community physicians and de-identified claims data for patients of those physicians community-wide comparison we expected limited power to detect differences in the measured numerators and denominator, especially without the permission to identify cohorts of patients eligible to receive chronic care management for chronic conditions. While we believed a feasibility study would be of interest, it would have been difficult to justify the data requests to health plans and the QDC given the small sample size and the consequently severely limited power of analyses (especially for the chronic disease cohorts). We requested a third no-cost extension to complete the medical record review fieldwork and obtain administrative data from health plans, but the request was not approved.

Inclusion of AHRQ Priority Populations

Our project focused on three Massachusetts eHealth Collaborative pilot communities. The pilot communities serve several AHRQ inclusion populations, specifically low-income and minority populations in one urban community and rural populations in a separate community.

Related studies of the MAeHC practices showed that they resembled practices throughout Massachusetts in terms of practice size, physician age and sex, prevailing financial incentives for quality performance and HIT adoption, and available resources for practice expansion. MAeHC practices were more likely to be located in rural areas (9.5% vs. 4.4%).

Our survey population was healthier and more affluent than the general population of Massachusetts. Among respondents to our survey, 97.1% reported that they were white, 0.7% black, 1.1% asian, 1.6% native american and 1.8% reported other race. Just over 1% reported hispanic family background. Two percent reported not graduating from high school. Slightly more than 8% reported fair or poor health status. Future evaluation projects may mitigate against these biases by engaging non-commercial insurance programs (Medicare and Medicaid) to recruit participants or recruiting patients directly through the health information exchange (as we had intended originally) or through clinical practices.

References

1. Schneider EC, Riehl V, Courte-Wienecke S, Eddy DM, Sennett C. Enhancing performance measurement: NCQA's road map for a health information framework. National Committee for Quality Assurance. *JAMA.* Sep 22-29 1999;282(12):1184-1190.
2. Schneider EC, Nadel MR, Zaslavsky AM, McGlynn EA. Assessment of the scientific soundness of clinical performance measures: a field test of the National Committee for Quality Assurance's colorectal cancer screening measure. *Arch Intern Med.* Apr 28 2008;168(8):876-882.

List of Publications and Products

Manuscript based on this report under development.