

Automating Quality Measurement: A System for Scalable, Comprehensive, and Routine Care Quality Assessment

Brian Hazlehurst, PhD¹, MaryAnn McBurnie, PhD¹, Richard Mularski, MD, MSHS, MCR¹,
Jon Puro, MPA-HA², Susan Chauvie, RN, MPA-HA²

¹Center for Health Research, Kaiser Permanente Northwest, Portland, OR

²OCHIN Inc, Portland, OR

Abstract

Electronic medical records (EMRs) hold the promise of making routine comprehensive measurement of care quality a reality. However, there are many informatics challenges that stand in the way of this goal. Guidelines are rarely stated in precise enough language for automated measurement of clinical practices and the data necessary for that measurement often reside in the text notes of EMRs. We designed a technology platform for scalable and routine measurement of care quality using comprehensive EMR data, including providers' free-text notes documenting clinical encounters. We are in the process of implementing this system to assess the quality of ambulatory asthma care in two diverse healthcare systems: a mid-size HMO and a consortium of Federally Qualified Healthcare Center (FQHC) clinics on the west coast of the United States.

Background

Comprehensive, routine, and meaningful quality measurement is needed to effectively guide quality improvement and care innovation. Unfortunately, comprehensive quality measurement, even within a selected domain of care delivery, is far from a routine process and involves expensive manual reviews of patient charts [1]. Electronic medical records promise to make routine comprehensive measurement of care quality a reality [2]. There are many informatics challenges, however, that stand in the way of realizing this objective. Primary among these are: (1) guidelines are not typically specified in a way that translates easily to quality measurement or computer implementation; (2) structured/coded data needed for quality measures are not standardized and are subject to local variations in EMR implementations and clinical practice; (3) much of the data required for quality measurement reside in free-text notes documenting clinical encounters.

We were funded by AHRQ to design and implement an automated and scalable method for comprehensive assessment of outpatient asthma care. We aimed to develop a quality measurement platform allowing for

the development of routine, up-to-date, comprehensive, and automated assessment of care quality, regardless of the targeted clinical domain.

Quality Measurement Methodology

Developing the measure set

Quality measurement draws from evidence-based recommendations for care, which are often summarized in clinical guidelines that target a particular domain of clinical care. Our study focused on outpatient care for asthma. To enable automated care assessment, recommended care steps delineated by guidelines must be converted into discrete quantifiable "measures." The development of each measure begins with a concise proposition about the recommended care for specific subsets of patients (e.g., "patients seen for asthma exacerbation should have a chest exam"). A comprehensive set of performance measures for a particular domain, once operationalized, constitute a "measure set." To ensure our measures would be comprehensive and current, we used an eight-stage iterative process to refine identified quality measures for ambulatory asthma care. This process included four separate vetting steps with local and national experts, soliciting independent comment and critique. We identified a starting set of 25 measures from comprehensive, rigorous, process quality measure sets, relying heavily on the RAND's Quality Assessment system [1,3,4] asthma measures. We added six proposed measures from recently revised guidelines [5] and other quality measurement sources (including HEDIS, NCQA, AMA, HRSA, and JACHO). We eliminated 10 measures that were not applicable to ambulatory care (n=6) or were inconsistent with current guidelines (n=4), resulting in a comprehensive set of 22 process measures.

Operationalizing the measure set

Next, each measure in the measure set must be converted into specifications of the component clinical events that can address two key properties: applicability of the measure to the patient (i.e., measure's denominator criteria met) and satisfaction of the measure by the provider's actions (i.e., numerator specifics met). Performance on each

measure across a population can then be reported as the percentage of patients who received recommended care as operationalized by the numerator criteria from among those for whom that care was indicated by meeting the denominator criteria. For example, the national RAND study of McGlynn and colleagues demonstrated that across 30 disease states, Americans received about 55% of recommended care [1].

For such a measurement scheme to be comprehensive, meaningful, and affordable, it requires that the necessary clinical events for each measure are routinely available in EMR data and can be extracted from the data warehouse. Thus, we investigated providers' clinical practices related to each measure, and also how that care is captured (through documentation) as data elements in the clinical information system and ultimately in the data warehouse. From these learnings, we have developed criteria for defining inclusion/exclusion in the denominator and numerator of each measure. Each measure's numerator requires a "measure interval," which is defined as the time window during which the recommended care delivery events are located for inclusion in the numerator. The measure interval is a time window oriented around some "index date" that is, in turn, a property of denominator inclusion. For example, for the measure that reads "patients seen for asthma exacerbation should have a chest exam," the index date is the exacerbation encounter, and the measure interval includes only that same encounter. On the other hand, the measure that reads "patients with persistent asthma should have a flu vaccination annually," the index date is the event that qualifies the patient as having persistent asthma and the measure interval is operationalized to include encounters six months prior to, through 12 months following, the index date. All patients included are guaranteed to have the same measure intervals and thus the same opportunity for receiving the recommended care. Table 1 shows a subset of the measures we developed for our outpatient asthma care quality (ACQ) study.

Applying the measure set

For a given measurement study, we first define an observation period (in our case, three years of clinical events captured in the EMR), and divide this into a period for denominator qualification (the "selection period") followed by a "evaluation period," during which, in most cases, the occurrences of prescribed care delivery are identified. In fact, each measure defines its own specific time intervals for qualification and evaluation, so this global division of the entire observation period provides only a general

picture of how the three years of clinical events are partitioned and included in measurement. We used a two-year selection period as an upper bound of time for identifying patients with "persistent asthma" (used in all of the measures in our set) or presenting at an office visit with an "asthma exacerbation" (used in 36% of the measures in our set). Persistent asthma patients were identified based on meeting minimum criteria for asthma-related utilization (e.g., medication orders or dispensings, outpatient visits or ED or hospital admissions). Patients also qualified for persistent asthma status through determination by the provider, as documented in her clinical notes. Asthma exacerbation criteria were based on hospitalization ICD-9 codes or an outpatient visit associated with a glucocorticoid order/dispensing and a text note indicative of exacerbation.

As mentioned above, measurement consists of assessing, for each measure, the count of patients who qualify for the measure and how many received the recommended care prescribed by the measure. All clinical events defining patient inclusion in the denominator must occur in the qualification period. Events defining inclusion in the numerator will most often occur within the evaluation period. In all cases, the clinical events included in the measurement study are limited to the observation period.

The ratios generated for each measure (counts of patients receiving the recommended care divided by counts of patients needing that care) can be produced at the patient, provider, clinic, and health-system levels. The key to scalable automation permitting this type of routine measurement is the reliable, maintainable, and comprehensive generation of the required clinical events, as defined by the measure set.

System Design

We designed a quality measurement system as a "pipeline" of transformation and markup steps taken on encounter-level electronic medical record data with the goal of capturing all of the clinical events required to assess care for specific clinical domains (see Figure 1). We are currently implementing this platform to assess outpatient asthma care in two distinct health systems: a mid-sized HMO and a consortium of Federally Qualified Healthcare Center (FQHC) clinics serving the uninsured and indigent in the west coast region. As shown in Figure 1, our system's pipeline can be divided into three sequential segments involving Data Extraction, Concept Markup, and Quality Measurement.

Quality Measure	Denominator criteria [Index Date]	Numerator criteria [Measure Interval]	Operationalization Comments
Patients with the diagnosis of persistent asthma should have a historical evaluation of asthma precipitants	Patients with persistent asthma [Qualification Date]	Patients with a subjective evaluation of precipitants listed in provider's notes [any documentation]	Probably only found in the text progress notes
Patients with the diagnosis of persistent asthma should have spirometry performed annually	Patients with persistent asthma [Qualification Date]	Patients with orders for PFTs or documentation of office spirometry [subsequent 12 months]	Numerator satisfied with documentation of referral to allergy or pulmonary specialist if no PFT known available with closed charting loop
Patients with the diagnosis of persistent asthma should have available short acting beta2-agonist inhaler for symptomatic relief of exacerbations	Patients with persistent asthma [Qualification Date]	Prescription for a short acting beta-2 agonist to use PRN [subsequent 12 months]	Numerator satisfied if prior / existing active Rx; also Ach or combination Rx (i.e. Combivent) or oral/nebulized PRN Rx will count. Exclusion if documented adverse reaction to β -agonists as allergy
Patients with persistent asthma should not receive non-cardioselective beta-blocker medications	Patients with persistent asthma [Qualification Date]	Pharmacy records without non-cardioselective beta-blocker prescription [subsequent 12 months]	e.g., nadolol, propranolol, pindolol
Patients with persistent asthma should have a flu vaccination annually	Patients with persistent asthma [Qualification Date]	Documentation of flu vaccination [prior 6 months or subsequent 12 months]	Numerator satisfied if vaccine documented regardless of where administered. Exclusion if documented egg allergy or patient refusal
All patients seen for an acute asthma exacerbation should have current medications reviewed	Patients with persistent asthma meeting criteria for outpatient exacerbation [exacerbation encounter]	Documentation that medications reviewed by provider [same visit]	Numerator satisfied if provider documents asthma specific medication history in notes or active management of current med list

Table 1. A sample of measures from the Asthma Care Quality (ACQ) measure set. (NOTE: Finalized measure set is available from authors upon request.)

Data Extraction

The data pipeline begins with extracts from the data warehouse of each EMR system. These extracts are produced by a component called the “EMR Adapter,” and contain the data required by the study, captured at the clinical-encounter level for all patients in the study population. In our study, this included the coded diagnoses, problems, and medical history updates generated at the visit; the medications ordered, dispensed and noted as current or discontinued for the patient; the immunizations, allergies, and health maintenance topics addressed at

the visit; as well as procedures ordered and progress notes and patient instructions generated for the visit. The population for our measurement study included all patients who had at least one asthma related visit (i.e., an asthma diagnosis code applied to the visit) during the three-year observation period.

The data are exported from the EMR data warehouse (typically, a relational database) into file-based eXtensible Markup Language (XML) documents according to a specification that is local to each data environment. The first transformation step in the

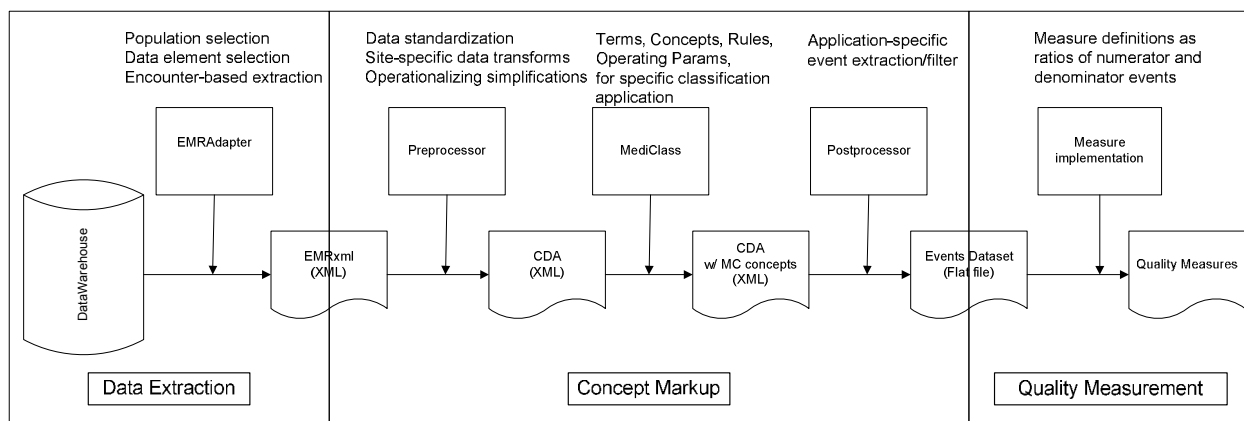


Figure 1. A system for automated quality measurement, shown as a pipeline of data transformation steps.

pipeline involves converting these locally-defined XML formats into a common, standard XML format conforming to the HL7 CDA specification for encounter data [6]. An XSLT [7] program written specifically for each data environment accomplishes this translation. We anticipate that EMR vendors will soon make available facilities for extracting CDA formatted encounter data directly from their EMR systems, potentially rendering this step unnecessary. However, whether these facilities will include the flexibility required to define the wide range of data needed for research purposes remains to be seen.

Concept Markup

The CDA provides a canonical representation of encounter-level data that is used as an input to our medical record classification system called MediClass [8]. MediClass uses natural language processing and rules defining logical combinations of marked up and originally coded data to generate concepts that are then inserted into the CDA document and passed along to the next step. This system has been successfully used to assess guideline adherence for smoking cessation care [9], to identify adverse events due to vaccines [10], and other applications that require extracting specific clinical data from text notes of the EMR. In the ACQ measure set, 46% of the measures require—and another 27% are enhanced by—processing the providers’ text notations to generate measure numerator events. In addition, all of the measure denominators in the ACQ measure set include criteria that are found in providers’ text notations.

Up to this point in the sequence, data processing is performed on-site within the secure data environments of each study site. This arrangement permits local control of sensitive data that resides in text notes and also in the comprehensive encounter record captured in CDA format. The next step filters these data to identify only those clinical events (including specific concepts identified in the text notes) that participate in the quality measures of the study. This step uses an XSLT program to process the marked-up CDA documents to produce a single file of measure-set specific clinical event data in comma-delimited format. This file is called the “Events Dataset.” Each line in this file identifies the study-coded patient, provider, and encounter, along with a single “event” (and attributes specific to that event) that participates in one or more measures of the measure set. Table 2 shows a partial schema for this events dataset file. This file (with IRB approval and a Data Use Agreement executed between the respective research organizations) is transferred from the multiple study sites to a central analysis location for final processing.

Quality Measurement

The distinct pipelines located at each health system converge into a single analysis environment for computation of quality measures. Here, information contained in the events dataset is processed across events to provide the clinical and temporal criteria for identifying patients that meet numerator and denominator criteria for each measure. Finally, the proportion of patients receiving recommended services is computed at the desired level (e.g., patient, provider, or health care organization).

EventType	Source Data Element	Eval1	Eval2	Eval3	Estr1	Estr2	Estr3	Estr4
MedsDisp	MedicationDispense	Daily Dose	Qty	# Refills	NDC	SIG	Name	Route
Asthma-Visit	VisitDx, MedicalHx, ProblemList, DischargeDx, ProgressNote	493.* ICD9 code			Asthma notations	Exacerbation notations		
Precipitant	ProgressNote, MedicalHx_comment, ProblemList_comment				Asthma Precipitant notations			
Spiro	Order, Referral, PFT_Result, MedicalHx_comment, ProblemList_comment, ProgressNote				Any spirometry notations			

Table 2. A portion of the schema defining Asthma Care Quality events in the (comma-delimited flat file) called the clinical events dataset. Each event type is derived from specific source data and has up to 7 data fields (four numerical values and four string values) for defining attributes associated with the event. Each event record is accompanied by data on patient, provider, date and location of care, etc.

Conclusion

Comprehensive and routine quality of care assessment requires not only state-of-the-art EMR implementations, but also a scalable technology platform for reliable, maintainable, and automated measurement of complex clinical practices. We have designed a system with the associated informatics challenges in mind and are currently implementing this system in two diverse healthcare systems to assess outpatient delivery of care to asthma patients. Our design overcomes challenges created by text-based guidelines, non-standard data elements, and text clinical notes that contain a large percentage of relevant data for measuring care quality. It remains to be seen if our design will accommodate all of the measures in our measure set, and if our implementations of the design produce accurate and meaningful measurement of the complex clinical practices in these diverse healthcare settings. It also remains to be seen if this system can, as intended, scale to address any target clinical domain of interest. Answers to these important questions are currently being addressed with ongoing research efforts.

References

1. McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A, Kerr EA. The quality of health care delivered to adults in the United States. *N Engl J Med* 2003; 348(26):2635-45.
2. Corrigan J, Donaldson MS, Kohn LT, (Eds). *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: National Academy Press; 2001.
3. Mularski RA, Asch SM, Shrank WH, Kerr EA, Setodji C, Adams J, Keesey J, McGlynn EA. The quality of obstructive lung disease care for adults in the United States: Adherence to recommended processes. *Chest* 2006 Dec; 130(6):1844-1850.
4. Kerr EA, Asch SM, Hamilton EG, McGlynn, EA. Quality of care for cardiopulmonary conditions: a review of the literature and quality indicators. Santa Monica: RAND; 2000.
5. National Asthma Education and Prevention Program, National Heart, Lung, and Blood Institute. Expert panel report 3: Guidelines for the diagnosis and management of asthma. Bethesda, MD: National Institutes of Health; 2007.
6. Dolin RH, Alschuler L, Beebe C, Biron PV, Boyer SL, Essin D, Kimber E, Lincoln T, Mattison JE. The HL7 Clinical Document Architecture. *J Am Med Inform Assoc*. 2001 Nov-Dec;8(6):552-69.
7. <http://www.w3.org/TR/xslt>
8. Hazlehurst B, Frost HR, Sittig DF, Stevens VJ. MediClass: A system for detecting and classifying encounter-based clinical events in any electronic medical record. *J Am Med Inform Assoc* 2005 Sep; 12(5):517-529.
9. Hazlehurst B, Sittig DF, Stevens VJ, Smith KS, Hollis JF, Vogt TM, et al. Natural language processing in the electronic medical record: assessing clinician adherence to tobacco treatment guidelines. *Am J Prev Med* 2005 Dec; 29(5):434-439.
10. Hazlehurst B, Naleway A, Mullooly J. Detecting possible vaccine adverse events in clinical notes of the electronic medical record. *Vaccine* 2009 Mar 23; 27(14):2077-2083.