

New paradigms for measuring clinical performance using electronic health records

JONATHAN P. WEINER^{1,2}, JINNET B. FOWLES³ AND KITTY S. CHAN¹

¹Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA, ²Division of Health Sciences Informatics, Johns Hopkins University, Baltimore, MD, USA, and ³Formerly with the Health Research Center, Park Nicollet Institute, Minneapolis, MN, USA

Address reprint requests to: Jonathan P. Weiner, 624 North Broadway, Room 605, Baltimore, MD 21205 USA. Tel: +1-410-955-5661; Fax: +1-410-955-0470; E-mail: jweiner@jhsph.edu

Accepted for publication 3 February 2012

Abstract

Measures of provider success are the centerpiece of quality improvement and pay-for-performance programs around the globe. In most nations, these measures are derived from administrative records, paper charts and consumer surveys; increasingly, electronic patient record systems are also being used. We use the term '*e-QMs*' to describe quality measures that are based on data found within electronic health records and other related health information technology (HIT). We offer a framework or typology for *e-QMs* and describe opportunities and impediments associated with the transition from old to new data sources. If public and private systems of care are to effectively use HIT to support and evaluate health-care system quality and safety, the quality measurement field must embrace new paradigms and strategically address a series of technical, conceptual and practical challenges.

Keywords: quality of care, electronic health records, health information technology, quality measurement, performance measurement, meaningful use

Measuring quality in the era of electronic health records

Few issues will be more important to the future of health care worldwide than the adoption and implementation of electronic health records (EHRs). Breakthrough technologies require shifts in paradigms. This shift has not yet occurred on a wide scale with regard to the use of EHRs for measuring and monitoring clinician and system performance.

EHRs and related health information technology (HIT), such as computerized provider order entry (CPOE), clinical decision support systems (CDS) and web-based personal health records (PHRs), have been cited as the foundation of the bridge across the health-care system's 'quality chasm' [1, 2]. In the USA, the Obama Administration's so-called *HITECH* program represents one of the largest HIT infrastructure investments a nation has ever made. It also represents the electronic framework that will underpin US health reform which will unfold over the next few years [3]. Moreover, similar scenarios are playing out in most other nations, where HIT systems are being built, expanded or inter-connected across isolated providers [4, 5].

The goal of all this is to improve workflow and documentation to help make care safer and more efficient. Though as

this is accomplished, we also need to be on the lookout for potential inadvertent harm that can be caused by suboptimal use of HIT systems [6]. Of considerable relevance to the quality improvement (QI) community is the benefit of EHRs as a tool for evaluating and improving clinician and system performance. Central to such QI applications are the measures that serve as the markers of success or failure. Moreover, these same measures are now frequently linked to provider financial incentives as part of 'pay-for-performance' (P4P) programs [7].

We use the term *electronic quality measures*, or *e-QMs*, to describe EHR-based performance measures. The implications of the shift toward electronic measures of quality are manifold. New ways of thinking are needed at this time, so we can integrate QI and pay for performance incentive programs into rapidly evolving EHR systems as they are being developed rather than after the fact; a much more difficult task.

The objectives of this article are: (i) to offer a typology of electronic measures of quality and safety; and (ii) to identify key challenges, opportunities and future priorities related to the development and application of these measures. Our work emerged from a project involving the development and application of *e-QMs* in the ambulatory care environment

[8, 9]. While our discussion draws heavily on experiences in the USA, we also obtained related information from other nations with well-developed primary and secondary care EHRs. Therefore, our insights should be relevant to high-to-middle income nations around the globe where similar issues are being confronted.

Though e-QMs are expected to become the norm in the USA in the not-too-distant future, that is not now the case. While over half of American doctor offices have some component of an EHR system, it is estimated that <25% of US ambulatory care is substantially documented by EHRs and fewer than 10% of these systems are both comprehensive and interoperable across providers [10]. Today, for most US providers, performance can be documented only with computerized insurance claims, abstractions of a limited sample of paper charts or surveys of a small subset of consumers. In other nations where EHRs are more widely established within primary care, EHR-based quality reporting is not uncommon. But given the limited cross-provider interoperability and lack of full integration of advanced HIT components, the e-QM capabilities in most other nations is still in early stages of development as well [11].

A framework for electronic quality measures

We developed a typology to help understand and shape the evolving field of electronic quality measurement. Our premise is that new ways of thinking about quality measurement are needed at this time, given the many unique functions and capabilities of EHRs and other HIT. Our e-QM framework (which is an extension of one we presented previously in a web-based report [8]) is meant to complement rather than displace other accepted quality constructs. We hope this classification exercise will support discourse between informaticians and EHR developers, QI specialists, clinicians, managers and policymakers.

The first three of our proposed five electronic quality measurement categories represent measures of care achieved by the provider. They reflect increasing degrees of reliance on unique EHR/HIT attributes. For that reason, we label them 'Level 1 to Level 3' e-QMs. The fourth and fifth measures focus on success of the EHR/HIT system implementation and unintended consequences, respectively. Our typology is outlined in Table 1 and discussed below.

Translated e-QMs

These are measures adapted from existing 'traditional' (i.e. not EHR supported) measurement sets. In the USA, 'traditional' refers to widely adopted measures such as those of the National Committee for Quality Assurance (NCQA) [12], which were originally designed for non-EHR data sources, such as paper medical charts and insurance claims. Outside the USA, claims data are less common, but many

Table 1 HIT/EHR-based e-quality measures (e-QMs) of provider performance

1. Translated (Level-1 e-QM): Measures based on traditional data collection approaches such as administrative (e.g. insurance claims) records which have been translated for use with EHR-based platforms (example: process of care measure such as percentage in scope patients receiving a lab test or mammography screening)
2. HIT assisted (Level-2 e-QM): Measures that while possible with non-EHR data sources would not be operationally feasible without the assistance of HIT (examples: blood pressure or BMI information on 100% of target population)
3. HIT enabled (Level-3 e-QM): Innovative measures that are not possible without comprehensive HIT (examples: percentage of abnormal test results read and acted upon by a clinician within 24 h, percentage of in-scope patient encounters where decision support modules were appropriately applied)
4. HIT system management: Measures used primarily to manage and evaluate HIT systems. (Examples: percentage of all prescriptions ordered via e-prescribing, percentage of EHR 'front sections' that are updated periodically.) These can be considered a measure of care structure
5. e-iatrogenesis: Measures of patient harm caused at least in part by the HIT system. (Example: the percentage of patients for whom the wrong drug was ordered due to an error intrinsic to an e-prescribing system; percentage of critical lab findings that did not result in patient notification)

HIT, health information technology; EHR, electronic health record.

This table is based in part on a previous publication [8].

settings do rely on data derived from non-EHR electronic data systems such as primary care clinic or hospital outpatient or inpatient administrative reporting systems.

Examples of translated e-QMs would include:

- The number of patients with diabetes seeing an ophthalmologist for an eye-care exam during a year.
- The number of children receiving appropriate immunizations.
- The number of women who have received a mammography within a given time frame.

From an informatics perspective, we consider these e-QMs to be the most basic 'Level 1'. That is, given their origin, such translated measures do not take advantage of any unique capabilities of EHRs. Level-1 e-QMs are ubiquitous in American integrated delivery systems because such measures are generally required by external agencies (e.g. federal, state and private payers) for all providers whether or not they have EHRs. Issues that surround the comparability of EHR-based measures with those derived traditionally (e.g. from claims or paper chart abstracts) have been the subject of a number of studies [13, 14].

HIT-assisted e-QMs

These are measures that, while not conceptually limited to EHR systems, would not be operationally feasible in settings without advanced HIT platforms.

Examples of HIT-assisted measure include:

- clinical outcomes of 100% of a patient panel based on physiologic measures such as body mass index or blood pressure;
- results of history and physical, or laboratory tests;
- percentage of prescriptions for a specific drug category (e.g. a β -blocker) written by the provider within a target time frame.

These measures are ‘Level 2’ because, while they do not take full advantage of the advanced properties of HIT systems, they do make use of the EHR’s ability to capture and integrate medical information not generally found in administrative sources. While most organizations could theoretically replicate these measures using manual abstraction of paper-based charts, it would be impractical to do so.

A recent article involving a consensus panel of US experts suggested a number of HIT-assisted e-QMs [15]. Level-2 HIT-assisted e-QMs are now common in wired US integrated delivery systems and in many international settings. In these environments, automated chart abstraction has supplanted manual chart reviews for QI measure reporting. For example, most measures used by the UK’s high profile ‘Quality Outcome Framework’ fall into this Level-2 e-QM category [16]. That is, they are based largely on information found within a GP practice site’s electronic patient record, though they could be derived from paper charts if needed.

HIT-enabled e-QMs

These are innovative e-measures that would not be possible outside of the HIT context. These are ‘Level-3’ measures because they embrace one or more unique HIT capabilities not readily found in paper charts. One such dimension is the ‘time-stamp’ capability of EHRs and CPOE systems where it is possible to know when the clinician received, viewed or acted on a specific item of information. Another unique capability of HIT is full information integration (sometimes termed interoperability) across all providers in a region. Other advanced HIT functions that are ripe for Level-3 e-QM development are those that go beyond charting. These would include: order entry systems (e.g. e-prescribing) and decision support systems (that recommend a course of action to the clinician); networked biometric devices (that capture the patient’s real-time physiologic function) and interactive web-based ‘personal health record’ systems (that capture the consumers preferences, functional status or satisfaction with care).

Some examples of HIT-enabled Level-3 e-QMs that could only be implemented in digitally supported settings include:

- Percent of clinicians reviewing new items in a chart (e.g. an out-of-range lab result) within x hours, and

acting on critical information (e.g. contacting the patient or ordering a follow-up test) within y hours of the instant they became aware of the event;

- Percentage of consultants who electronically shared specific pieces of information with the patient’s primary care doctor within x days of seeing the patient;
- Percentage of in-scope patients for whom real-time clinical decision support module have been appropriately applied (and ‘alerts’ were not ignored);
- Percentage of in-scope consumers who used their home monitoring device to digitally notify the provider of a reportable event (e.g. high blood sugar or high blood pressure) and the percentage of instances where HIT-mediated follow-up occurred.

These types of Level-3 e-QMs represent the frontier of quality measurement development, but they are not yet common; in larger part, because few providers’ EHRs are fully interoperable with all others in a region and few have fully integrated decision support, order entry and linked consumer ‘web portals.’ Moreover, even when these digital attributes are in place, there have been limited attempts by government and other regulators to move mandated performance measures beyond Level 1 or 2 (i.e. those feasible with administrative records, paper charts or stand-alone basic electronic charts). The reason for this is the least common denominator concern; until most providers have advanced HIT capabilities, the less advanced sites would be left out of the reporting systems.

Among advanced integrated delivery systems in the USA, such indicators are being piloted, but they are not yet used on a wide scale. Outside of the USA, even in nations with advanced quality performance indicator frameworks and high EHR penetration (e.g. the UK and Sweden, Denmark, New Zealand), Level-3 measures that embrace the unique capabilities of HIT systems have not been widely adopted [11, 16, 17].

HIT-system-management e-QMs

These measures are needed to support the deployment, management, evaluation and improvement of HIT systems. They can be used by organizations implementing the EHR or an external body wishing to evaluate a provider’s HIT system.

Given that many believe that an HIT system is an essential requirement for a health delivery system’s infrastructure in the twenty-first century, these can be considered a type of ‘structure’ measure, following the Donabedian structure/process/outcome framework.

Examples of HIT-system-management measures include:

- EHR item-completion rates;
- attainment of community interoperability targets;
- presence of various computerized decision support algorithms;
- percentage of real-time CDS alerts bypassed by clinicians;

- percentage of patient-allergy lists reviewed by patients (e.g. via a web portal system where patients can view their own record) annually;
- proportion of key variables lost to measurement due to use of non-standard free-text EHR notation which cannot be accessed.

As part of the US Federal Department of Health and Human Services' EHR expansion initiative, the Center for Medicare/Medicaid Services (CMS) and Office of the National Coordinator for HIT (ONC) are offering incentive payments linked to what they term the 'meaningful use' of EHRs [3, 18, 19]. The sums are large: up to (US)\$44 000 per doctor and more than \$2 million per hospital over 3 years (2011–13). During the first stage of this program (started in 2011), providers must achieve success on measures that largely fall into this systems management category; for example, maintaining certain key sections in their EHR and being able to exchange information with other providers [19].

'e-iatrogenesis' e-QMs

These measures document patient harm caused at least in part by the application of HIT [6]. Specifically, such measures assess the degree to which unanticipated quality and safety problems arise through the use of HIT, whether the error is of human (provider or patient), technological or organizational/system origin. These problems may involve errors of commission or omission [20–22].

Examples of e-iatrogenesis targeted e-QMs include:

- Percentage of patients receiving incorrect medications or procedures because of HIT-related errors in the CPOE/e-prescribing process;
- Number of sentinel event computerized decision support errors (either of 'omission' or 'commission') impacting patients;
- Number of critical provider or patient e-notifications not received resulting in patient harm.

e-iatrogenesis and other types of unanticipated harms and errors linked to HIT are expected to increase significantly in coming years. Both informatics and quality experts agree that this domain should be a high priority of safety and QI programs. Identifying measures to support these monitoring programs will be essential [23, 24].

Digital opportunities: digital challenges

As the availability of electronic data expands exponentially, the technical, conceptual and practical challenges that must be confronted by developers of health-care performance measures will increase as well.

Clinician usability is the key to successful EHR adoption. Clinicians are not willing to click through endless menus or 'pop up alerts', preferring free-text notation as is the norm for paper charts. While this situation is well known to EHR

designers [24, 25], the implications for those developing e-QMs are both profound and not well understood.

The current key source of performance measures in the USA is insurance claims data, which are a byproduct of the financial transaction process. Accordingly, measures based on claims items with high relevance for the payment process are more likely to be reliable and valid [26–28]. In the HIT age, this dynamic will change; rather than payment, the accuracy axiom will revolve around the structured care workflow. Quality measures based on the unstructured fields within the EHR, e.g. free-text charting, will be far less accurate [29, 30].

As practices implement EHRs, they must also be careful not to take steps backwards as e-QM-based metrics replace traditional quality reporting. For example, unless structured electronic templates are fully (and accurately) accepted by clinicians using the EHRs, it may be necessary to manually abstract the EHR's free-text sections to assess whether process standards of care were met or certain patient outcomes were achieved. This problem will continue until required quality indicators are entered (either manually or via electronic streaming) in their appropriate structured location within the record or until vastly improved natural language processing systems—capable of 'mining' the clinician's free-text chart entries—are implemented.

There appears to be unrealistic expectations among clinical, management payer and government stakeholders regarding the ease with which EHRs can be used to derive quality measures. A widely held, but erroneous, belief is that once an EHR is implemented by a health-care delivery organization, just hitting the 'F9 key' will lead to an immediate and effortless flow of provider performance reports. Moving forward, these expectations must be more effectively managed.

More and better data should be the hallmark of well designed and implemented EHR systems. But the 'better' part of this 'more + better' equation cannot be taken for granted: automated data are not necessarily improved data. Serious unanticipated data quality problems have been reported in HIT systems that were poorly conceived or executed [8]. For example, keeping the EHR's 'front page' section (e.g. problems, allergies and medication lists), current is a major challenge. Without strict protocols for verifying, organizing, updating and purging information found in EHRs, these repositories can become unstructured 'electronic attics' of uncertain accuracy that will be of limited value to QI analysts.

The HIT-supported care process may also introduce measurement challenges that are not well understood. Clinicians increasingly will follow real-time 'care pathways', where the EHR mediates much of the doctor's diagnostic and therapeutic actions as well as their interaction with the patient. Furthermore, the care process will be self-documenting based on this EHR-directed workflow. From a measurement perspective, there will be a high degree of circularity between these 'hard-wired' digitally mediated actions and the e-QM criteria that are used to measure ideal performance. For example, if the doctor clicks on 'yes, I discussed all relevant issues with the patient' or yes, 'the patient should receive the recommended lab or drug order set', then

by default, the patient will meet the ideal standard of care as defined by an e-QM that is derived from the EHR data warehouse. This sequence of events suggests that first, we need to be cautious about the accuracy of these 'clicks' (e.g. did the doctor really discuss smoking with the patient or just click through the required screen) and second, in cases where the click translates to a specific diagnostic or therapeutic action (e.g. the ordering a standard battery of tests), it will be essential that the embedded algorithms be fully evidence-based. Otherwise, we risk promulgating ineffective care. The HIT, QI and comparative effectiveness research communities will need to partner closely in addressing this important potential problem.

HIT will change the time dimension of the QI process. Rather than annual or quarterly retrospective reports, new models of quality reporting will likely involve 'dashboard' indicator panels that can be monitored by both individual clinicians and managers in real time. We will need to learn more about the implications of this shift in temporal perspective on all key parties as well as the HIT and QI systems that are the source of these indicators.

The EHR and HIT systems that are rapidly diffusing around the globe hold great promise for improving health-care effectiveness. But today, even in advanced settings that are well supported by EHRs, quality performance measures and mindsets remain heavily anchored to constructs and contexts previously applied using paper charts and administrative records. As health-care delivery systems become digital, so too must the infrastructure we use to measure the performance of these systems. The electronic quality measures we develop for this purpose will need to be as innovative as the HIT systems themselves.

Acknowledgements

We are most grateful to the many organizations which provided us input regarding the application of EHRs to quality measurement within their organizations. We are especially grateful to members of a consortium of US integrated delivery systems that work with us on this project. These included: Billings Clinic, Billings, Montana; Geisinger Health System, Danville, Pennsylvania; HealthPartners, Minneapolis, Minnesota; Kaiser Permanente of the Northwest, Portland, Oregon; and Park Nicollet Health Services, Minneapolis, Minnesota. The assistance of Elizabeth Kind, Toni Kfuri, MD, and Jessica Holzer is gratefully acknowledged. The key informants at the integrated delivery systems who provided us valuable input included: Patricia Coon, MD, Mark Selna, MD, Leif Solberg, MD, Andrew Nelson, Lynne Dancha, Dean Sittig, PhD, Nancy Jarvis, MD and Shannon Neale, MD.

Funding

This work was funded with grants from the Commonwealth Fund (grant number 20060073) and the Robert Wood

Johnson Foundation (grant number 053729). Seed grant funding was also received from the Agency for Healthcare Quality and Research (grant number 275-JHU-01).

References

1. Corrigan J, McNeill D. Building organization capacity: a cornerstone of health system reform. *Health Affairs* 2009;**28**:w-205–15 (27 January 2009, date last accessed).
2. Walker JM, Carayon P. From task to process: the case for changing health information technology to improve health care. *Health Affairs* 2009;**28**:467–77.
3. Blumenthal D. Stimulating the adoption of health information technology. *N Engl J Med* 2009;**360**:1477–9.
4. Ovreteit J, Scott T, Rundall T *et al.* Improving quality through effective implementation of information technology in health-care. *Int J Qual Health Care* 2007;**19**:259–66.
5. Robertson A, Cresswell K, Takian A *et al.* Implementation and adoption of nationwide electronic health records in secondary care in England; qualitative analysis of interim results from a prospective national evaluation. *BMJ* 2010;**341**:c4564, online first.
6. Weiner JP, Kfuri T, Chan K *et al.* 'e-Iatrogenesis': the most critical unintended consequence of CPOE and other HIT. *J Am Med Inform Assoc* 2007;**14**:387–8.
7. Institute of Medicine. *Rewarding Provider Performance: Aligning Incentives in Medicare*. Washington, DC: National Academy of Sciences Press, 2007.
8. Fowles JB, Weiner JP, Chan K *et al.* *Performance Measures Using Electronic Health Records: Five Case Studies*. New York, NY: The Commonwealth Fund, 2008 [Internet]. http://www.commonwealthfund.org/usr_doc/Briggs_Fowles_perfmecasEHRs5casesudies_1132.pdf?section=4039 (25 November, 2011, date last accessed).
9. Chan K, Weiner J. EHR based quality indicators for ambulatory care: findings from a review of the literature. Narrative Report to AHRQ, 2006 [Internet]. http://healthit.ahrq.gov/portal/server.pt/gateway/PTARGS_0_3882_217665_0_0_18/e-indicator-lit-review.pdf (25 November, 2011, date last accessed).
10. Hsiao C, Hing E, Socey TC *et al.* Electronic medical record / electronic health record systems of office-based physicians: United States, 2009 and preliminary 2010 state estimates. National Center for Health Statistics, 2010 [Internet]. http://www.cdc.gov/nchs/data/hestat/emr_chr_09/emr_chr_09.pdf (25 November 2011, date last accessed).
11. Gray B, Bowden T, Johansen I *et al.* *Electronic Health Records: An International Perspective on 'Meaningful Use'*, Issues Brief, November 2011. New York: The Commonwealth Fund, 2011 [Internet]. (Analysis of Sweden, Denmark, New Zealand). <http://www.commonwealthfund.org/Publications/Issue-Briefs/2011/Nov/Electronic-Health-Records-International-Use.aspx> (25 November 2011, date last accessed).
12. National Committee for Quality Assurance. The Healthcare Effectiveness Data and Information Set (HEDIS) program, 2012 [Internet]. <http://www.ncqa.org/tabid/59/Default.aspx> (25 November 2011, date last accessed).

13. Tang PC, Ralston M, Arrigotti MF *et al.* Comparison of methodologies for calculating quality measures based on administrative data versus clinical data from an electronic health record system: implications for performance measures. *J Am Med Inform Assoc* 2007;**14**:10–5.
14. Linder JA, Kaleba EO, Kmetick KS. Using electronic health records to measure physician performance for acute conditions in primary care: empirical evaluation of the community-acquired pneumonia clinical quality measure set. *Med Care* 2009;**47**:208–16.
15. Kern LM, Dhopeswarkar R, Barron Y *et al.* Measuring the effects of health information technology on quality of care: a novel set of proposed metrics for electronic quality reporting. *Jt Comm J Qual Patient Saf* 2009;**35**:359–69.
16. Indicators for Quality Improvement. United Kingdom National Health Service Information Centre for Health and Social Care, 2010 [Internet]. <https://mqi.ic.nhs.uk/> (25 November 2011, date last accessed).
17. Socialstyrelsen and Swedish Association of Local Authorities and Regions. Quality and Efficiency in Swedish Health Care: Regional Comparisons, 2009 [Internet]. <http://www.socialstyrelsen.se/Lists/Artikelkatalog/Attachments/18023/2010-4-37.pdf> (25 November 2011, date last accessed).
18. Hogan S, Kissam S. Measuring meaningful use. *Health Affairs* 2010;**29**:601–6.
19. Blumenthal D, Tavenner M. The ‘meaningful use’ regulation for Electronic Health Records. *N Engl J Med* 2010;**363**:501–4.
20. Ash J, Sittig DF, Poon EG *et al.* The extent and importance of unintended consequences related to computerized provider order entry. *J Am Med Inform Assoc* 2007;**14**:415–23.
21. Metzger J, Welebob E, Bates DW *et al.* Mixed results in the safety performance of computerized physician order entry. *Health Affairs* 2010;**29**:655–63.
22. Koppel R, Metlay JP, Cohen A *et al.* Role of computerized physician order entry systems in facilitating medication errors. *JAMA* 2005;**293**:1197–203.
23. Bloomrosen M, Starren J, Lorenzi NM *et al.* Anticipating and addressing the unintended consequences of health IT and policy: a report from the AMIA 2009 Health Policy Meeting. *J Am Med Inform Assoc* 2011;**18**:82–90.
24. Institute of Medicine. *Health IT and Patient Safety: Building Safer Systems for Better Care*. Washington, DC: National Academies Press, 2011.
25. Berner ES, Moss J. Informatics challenges for the impending patient information explosion. *J Am Med Inform Assoc* 2005;**12**:614–7.
26. Weiner JP, Powe N, Steinwachs D *et al.* Applying insurance claims data to assess quality of care: a compilation of potential indicators. *Qual Rev Bull* 1990;**16**:424–38.
27. Fowles J, Lawthers A, Weiner J *et al.* Agreement between physicians’ office records and medicare claims data. *Health Care Financ Rev* 1995;**16**:189–99.
28. Iezzoni LI. Assessing quality using administrative data. *Ann Intern Med* 1997;**127**:666–74.
29. Bridges to Excellence Consortia. Measuring What Matters: Electronically, Automatically, (Somewhat) Painlessly, 2009 [Internet]. <http://www.rwjf.org/files/research/measuringwhatmatters2009.pdf> (23 March 2011, date last accessed).
30. Chan K, Fowles J, Weiner J. Electronic health records and the reliability and validity of quality measures: a review of the literature. *Med Care Res Rev* 2010;**67**:503–27.