

Performance Measures

Measuring the Effects of Health Information Technology on Quality of Care: A Novel Set of Proposed Metrics for Electronic Quality Reporting

Lisa M. Kern, M.D., M.P.H.; Rina Dhopeswarkar, M.P.H.; Yolanda Barrón, M.S.; Adam Wilcox, Ph.D.; Harold Pincus, M.D.; Rainu Kaushal, M.D., M.P.H.

Clinical information is frequently missing at the point of care, including in one out of every seven primary care visits.¹ The missing information is often clinically important, and its absence can result in harm to patients.^{1,2} Missing information also contributes to a highly fragmented health care system.³ As a top priority for national action, the Institute of Medicine recommends decreasing fragmentation through coordination of care, which requires exchange of information among participants responsible for different aspects of a patient's care.^{4,5}

Coordination of care is time-consuming and difficult. For example, to coordinate care for Medicare patients alone, the typical primary care physician must communicate with more than 200 physicians in more than 100 practices.⁶ In addition, primary care physicians require information from the many laboratories, pharmacies, radiology facilities, hospitals, health plans, and other organizations with which their patients interact.

Interoperable health information technology (HIT) is widely viewed as a potential solution, with its implementation being pursued by the federal government (most recently through the American Recovery and Reinvestment Act⁷) and three fourths of states.⁸ *HIT* refers broadly to electronic methods for storing and processing clinical data, whereas *interoperable* refers to the capacity to be linked electronically to other IT systems.⁹ Electronic health records (EHRs), one form of HIT, enable secure and reliable management of a patient's longitudinal medical information at the point of care.¹⁰ Health information exchange is another form of HIT, which enables electronic movement of health-related information among organizations.¹¹ Of note, the American Recovery and Reinvestment Act provides financial incentives for meaningful use of EHRs that support the exchange of data electronically.

No single metric set exists for measuring the effects of EHRs and health information exchange on quality of care. Existing quality metric sets do not presume communication between

Article-at-a-Glance

Background: Electronic health records (EHRs), in combination with health information exchange, are being promoted in the United States as a strategy for improving quality of care. No single metric set exists for measuring the effectiveness of these interventions. A set of quality metrics was sought that could be retrieved electronically and would be sensitive to the changes in quality that EHRs with health information exchange may contribute to ambulatory care.

Methods: A literature search identified quality metric sets for ambulatory care. Two rounds of quantitative rating of individual metrics were conducted. Metrics were developed de novo to capture additional expected effects of EHRs with health information exchange. A 36-member national expert panel validated the rating process and final metric set.

Results: Seventeen metric sets containing 1,064 individual metrics were identified; 510 metrics met inclusion criteria. Two rounds of rating narrowed these to 59 metrics and then to 18. The final 18 consisted of metrics for asthma, cardiovascular disease, congestive heart failure, diabetes, medication and allergy documentation, mental health, osteoporosis, and prevention. Fourteen metrics were developed de novo to address test ordering, medication management, referrals, follow-up after discharge, and revisits.

Discussion: The novel set of 32 metrics is proposed as suitable for electronic reporting to capture the potential quality effects of EHRs with health information exchange. This metric set may have broad utility as health information technology becomes increasingly common with funding from the federal stimulus package and other sources. This work may also stimulate discussion on improving how data are entered and extracted from clinically rich, electronic sources, with the goal of more accurately measuring and improving care.

health care providers and were not specifically designed to capture the potential incremental effect of receiving clinical data from external sources. Existing metric sets also largely assume that data would be derived from manual chart review, which is laborious and expensive, or claims data, which lack clinical detail. No existing metric set has been designed for electronic reporting from clinical sources (such as EHRs and/or health information exchanges), which could ultimately be less expensive than manual review and more clinically nuanced than claims.

Having a new metric set, which could be reported electronically and could capture any potential effects of EHRs with health information exchange, would have several applications. First, it could be used to measure the effectiveness of these technologies for improving coordination of care and, if found effective, spur adoption and inform iterative refinements. Second, it could be used to compare different electronic strategies for delivering clinical data to physicians—determining, for example, the relative effectiveness of feeding data directly into EHRs versus having freestanding Internet-based portals. This information would be informative, as the optimal strategy is unknown. Third, it could be used to isolate the effects of interoperable HIT in communities where other concurrent interventions, such as pay-for-performance or medical home transformation, are the primary subject of study. Finally, metric sets of this kind could encourage vendors to enhance capabilities for electronic reporting, an important refinement that would facilitate quantitative measurement of quality effects.

We therefore sought to develop a metric set that could be reported electronically and that could be sensitive to the types of improvements in quality that EHRs with health information exchange may contribute to the ambulatory setting.

Methods

Our methodology consisted of eight steps: (1) a literature review for existing ambulatory care quality metric sets; (2) application of exclusion criteria to individual metrics contained in the metric sets; (3) articulation of assumptions, a conceptual model, and domains for rating; (4) a first round of quantitative rating; (5) validation of the process by a national expert panel; (6) a second round of quantitative rating by national experts; (7) development of metrics de novo; and (8) validation of the final metrics by national experts.

1. LITERATURE REVIEW

We included metric sets that were endorsed by a national quality, professional, or research organization. We also specifi-

cally included metric sets that addressed quality of transitions across health care settings, regardless of whether they had been formally endorsed. We identified metric sets through five types of searches: Ovid Medline searches, Google searches, searches of the Web site of the Agency for Healthcare Research and Quality (<http://www.ahrq.gov>), review of the Institute of Medicine's *Performance Measurement*,¹² and expert recommendations.

2. APPLICATION OF EXCLUSION CRITERIA

We assembled the individual metrics from all metric sets retrieved and removed duplicates. To narrow the scope of the project, we excluded metrics that (1) were not relevant to ambulatory care (emergency department care was excluded); (2) were not relevant to adult primary care (obstetrics, pediatrics, cancer care, and HIV care were among the specialties excluded); (3) consisted of provider, practice, or health plan characteristics; or (4) captured patient or provider satisfaction.

3. ASSUMPTIONS, CONCEPTUAL MODEL, AND DOMAINS FOR RATING

We asked raters to assume the perspective of a board-certified, community-based primary care physician who has been in practice for 10 years, has a relatively stable panel of patients, and has an interoperable EHR that is currently linked only to generalist partners.

We developed a conceptual model (Figure 1, page 361), which assumed that some clinical data resided outside the physician's practice and that health information exchange would deliver those data electronically at or before the point of care. Any effect of health information exchange would then be mediated by the effect of the data on medical decision making (for example, ordering a prescription, ordering a test, referring to another provider, referring for hospitalization, or not performing any of those actions). For electronic reporting, the medical decision made (for example, prescription, test, referral, or hospitalization) would be exported electronically and aggregated across patients.

Following this model, we asked reviewers to rate each metric on two domains: "Sensitivity to the Potential Effects of EHRs plus Health Information Exchange" and "Suitability for Electronic Reporting." For the first domain, raters assigned each metric a score from 0 ("not at all sensitive") to 6 ("extremely sensitive") in response to the question "Overall, how sensitive is medical decision making to electronic receipt of the data elements that the health care provider would need to address the performance metric?" Raters inferred the types of data elements

Conceptual Framework for Developing an Electronically Reportable Metric Set

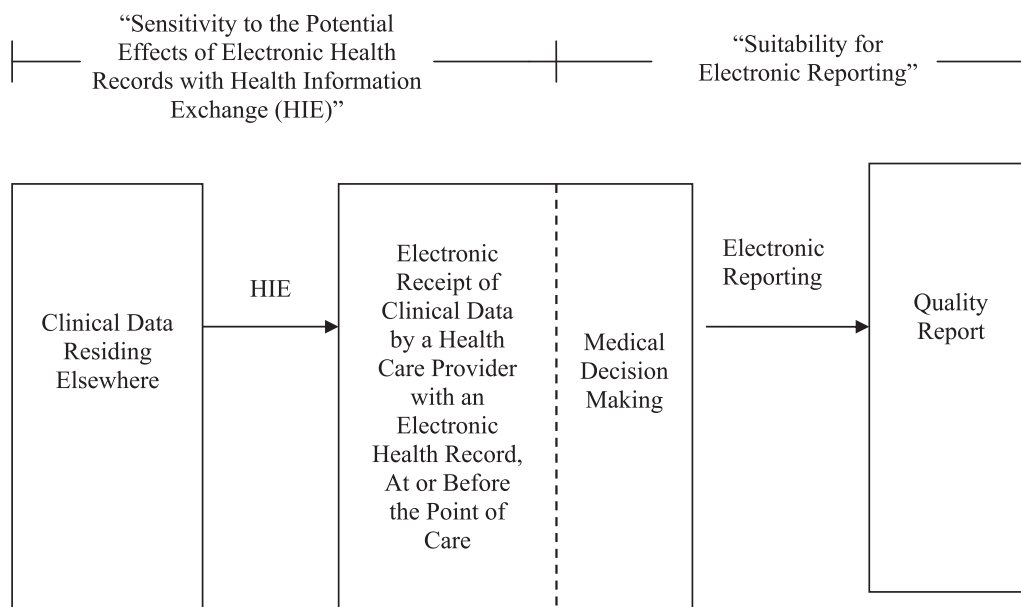


Figure 1. A conceptual framework for developing an electronically reportable metric set for capturing the potential effects of health information exchange, in the setting of electronic health records, is shown.

that health information exchange could provide that would be useful. For example, if the metric was use of beta-blockers in patients with a history of myocardial infarction, the rater assessed whether receiving diagnosis data, medication lists, and/or allergy lists would affect his or her prescribing decisions. We asked raters to consider implicitly the feasibility of delivering the data elements electronically.

For the second domain, raters assigned each metric a score from 0 (“not at all suitable”) to 6 (“extremely suitable”) in response to the question “Overall, how suitable is this metric for electronic reporting for quality measurement?” We asked raters to consider implicitly clinical importance, feasibility of electronic reporting, and validity of an electronically reported version of the metric.

4. ROUND ONE OF QUANTITATIVE RATING

We included seven raters in the first round of rating: three general internists [including L.M.K.], one internist-pediatrician [R.K.], one pediatrician, one psychiatrist [H.P.], and one informatician [A.W.], all of whom were also health services researchers.* Each rater was given at least two hours of project-

* The pediatrician was included for her expertise in health services research. However, given that pediatric metrics were excluded, the authors conducted a sensitivity analysis removing her ratings, and the results did not change.

specific training. Each metric was assigned to two raters. Raters were blinded to the organization that developed each metric. Within each rater, we added the scores for the two domains for each metric. We then averaged this summary score across raters. On the basis, in part, of the target number of metrics desired for the second round of rating, we considered metrics with summary scores ≥ 9 (on a scale from 0 to 12) to be high.

5. NATIONAL EXPERT PANEL

We convened a panel (Appendix 1, available in online article) that included 10 national experts in quality measurement and 5 national experts in interoperable HIT. We also included 7 leaders of community-based HIT initiatives funded by New York State, 4 representatives from the New York State Department of Health, and 4 from the New York eHealth Collaborative, as New York State has become a national leader in advancing community-based interoperable HIT.^{13,14} The panel was convened in person in Manhattan on April 7, 2008, and by phone on June 2, 2008, and June 23, 2008. In addition, we also consulted individually by phone with 6 other national experts who were unable to attend the other meetings. Of the 36 experts, 22 were also physicians (in addition to having the cited content expertise).

The national expert panel validated the process by which we

developed the metric set and confirmed the face validity of the high-scoring metrics from Round One. It also suggested several refinements for the next round of rating, which we incorporated.

6. ROUND TWO OF QUANTITATIVE RATING

Sixteen of the 36 expert panelists [including R.K. and H.P.], as well as two other original raters [L.M.K., A.W.] participated in Round Two, scoring each metric along five domains: feasibility of delivering data electronically to the physician at the point of care, impact on medical decision making, clinical importance, feasibility of reporting data electronically, and a global rating. We provided raters with detailed metric specifications, and we consistently used the wording endorsed by the National Quality Forum (NQF) when there were nearly identical metrics from different organizations.

Each metric was scored by 4 to 7 raters, and we averaged global scores across raters. Choosing scores above the scale's midpoint, we considered an average score of ≥ 4 (on a scale from 0 to 6) to be high. To consider variability among raters, we separately considered metrics that had both received high average scores and a low score (< 3) from at least one rater.

7. DEVELOPMENT OF METRICS DE NOVO

The first step in the development of de novo metrics was the identification of areas of quality expected to be affected by EHRs with health information exchange that were not represented by existing metrics. We drew in part on our extensive collective research experience working with communities that are implementing EHRs with health information exchange. We had studied the technologies that these communities were implementing and the types of data targeted for health information exchange.¹⁵ We also circulated to communities a comprehensive list of more than 100 HIT functionalities and elements of health information exchange for the ambulatory setting. Communities responded by indicating which functionalities they were planning to implement, as well as expected clinical and economic benefits. We then reviewed these responses, as well as suggestions from our expert panel, to identify areas for de novo metric development. We conducted literature reviews to determine previous work done in those areas, regardless of whether that work was reflected in existing metrics.

Following this process, we drafted a set of metrics and presented it to the national expert panel as a whole and to 10 national experts individually. This team of 10, many of whom had volunteered to participate in de novo metric development,

included experts in quality measurement, interoperable HIT, and community-based HIT initiatives. We iteratively refined the metrics with the experts to reach a final set. The process of refinement included adding metrics suggested by experts, combining metrics with similar content, removing metrics that were not suitable for electronic reporting, and rewording for clarity.

8. VALIDATION OF THE FINAL METRICS

The final metric set, including selected existing metrics and de novo metrics, was approved by the national expert panel.

Results

METRICS SELECTED FROM EXISTING METRIC SETS

We identified 17 metric sets for measuring ambulatory care quality (Table 1, page 363), which contained a total of 1,064 individual metrics. Of these, we identified 122 duplicates (Figure 2, page 363). We excluded 84 metrics not relevant to the ambulatory care setting; 136 not relevant to adult primary care; 189 consisting of provider, practice, or health plan characteristics; and 23 on patient or provider satisfaction.

The remaining 510 metrics underwent the first round of rating on two domains ("Sensitivity to the Potential Effects of EHRs plus Health Information Exchange" and "Suitability for Electronic Reporting"). For each domain, the distribution of scores (averaged across raters) was bimodal; a large proportion of metrics (approximately 20%) were rated less than 1 (not at all sensitive or not suitable for electronic reporting), and the rest had scores that were approximately normally distributed. The average summary score (combining the two domains) was 4.9 (standard deviation [SD], 2.8). Fifty-nine metrics received summary scores of 9 or more (out of 12).

For the second round of rating, global scores on the 59 metrics ranged from 0 to 6 within individual raters. When averaged across raters, the range of global scores narrowed to 2.6 to 5.0, with an overall mean of 3.8 (SD, 0.6). There were two groups of high-scoring metrics: those with mean global scores ≥ 4 and scores ≥ 3 from all raters (Group A, $n = 18$) and those with global scores ≥ 4 and a score < 3 from at least one rater (Group B, $n = 5$).

In discussions with the expert panel, one of the Group A metrics, which dealt with medication reconciliation, was removed from the final list because of concerns about feasibility of electronic reporting. The rest of the Group A metrics were chosen as the final set, plus one metric from Group B, regarding follow-up after discharge from a hospitalization for mental health, which the experts re-included due to the desire to have

Table 1. Existing Ambulatory Care Quality Metric Sets (N = 17)*

- The Ambulatory Care Quality Alliance (AQA) measures¹⁶
- The National Committee for Quality Assurance (NCQA) Healthcare Effectiveness and Data Information Set (HEDIS) measures¹⁷
- The National Quality Forum (NQF) ambulatory care measures¹⁸
- Take Care New York measures (extended set)¹⁹
- Subsets of the Assessing Care of Vulnerable Elders (ACOVE) measures:
Those applicable to both medical records and administrative data²⁰
Those cited by Eric Coleman as relevant for transitions across care settings²¹
- Doctor's Office Quality Information Technology (DOQIT) measures²²
- Physician Quality Reporting Initiative (PQRI) measures²³
- Care Transitions Measures (CTM-3)²¹
- Care Transitions Measures (CTM-15)²⁴
- Physician Practice Connections—Patient Centered Medical Home (PPC-PCMH) measures²⁵
- American College of Cardiology (ACC) and American Heart Association (AHA) chronic heart failure measures²⁶
- ACC/AHA/Physician Consortium for Performance Improvement (PCPI) chronic stable coronary artery disease measures²⁷
- American Diabetes Association/NCQA diabetes measures²⁸
- American Geriatrics Society proposed pay-for-performance measures²⁹
- Center for Quality Assessment and Improvement in Mental Health (CQAIMH) ambulatory measures³⁰
- Standards for Bipolar Excellence (StABLE) measures³¹
- Technical Appendix to McGlynn E.A. et al.: "Who Is at Greatest Risk for Receiving Poor-Quality Health Care?" part of the RAND Health Working Papers series.³²

* See References on page 369.

mental health represented in the final set.

The final 18 metrics are shown in Table 2 (page 364), along with the metric set from which each was drawn. The final set included 1 metric for asthma, 3 for cardiovascular disease, 1 for congestive heart failure, 4 for diabetes, 1 for mental health, 2 for medication and allergy management, 1 for osteoporosis, and 4 for prevention. The metrics included both processes of care (such as receiving a certain medication or test) and intermediate health outcomes (such as the degree of control of blood glucose or cholesterol).

LESSONS LEARNED FROM REVIEW AND SELECTION OF METRICS FROM EXISTING METRIC SETS

The face validity of the final metric set was supported by the

Flow Diagram Illustrating Selection of Existing Quality Metrics

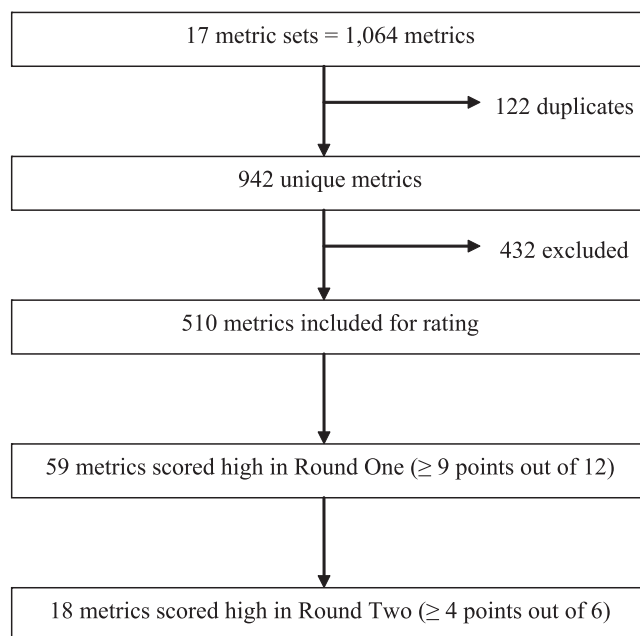


Figure 2. The process by which metrics were excluded from the total of 1,064 individual metrics is shown.

inclusion of metrics for diseases widely recognized as important public health problems. Also included in the final set were metrics on medication and allergy documentation, which were not disease-specific and not currently widely used. In discussions, experts reported that although documentation is only a part of quality measurement, medication and allergy management are important facets of care being targeted by EHRs with health information exchange. They further asserted that measurement of documentation would be a first step toward encouraging electronic reporting of, and perhaps even better performance on, more complex medication management metrics.

A common theme across many top-scoring metrics was the need for data from more than one source, such as diagnosis data plus medication data or diagnosis data plus laboratory data. This was mentioned by several experts as being intentional, representing the type of integration of data which one would expect from better coordination of care. One expert even argued that the ability of a community to report these metrics electronically should itself be an indicator of quality.

DE NOVO METRICS

We identified five topics not well represented among exist-

Table 2. Top-Scoring Existing Metrics for Measuring with Electronic Reporting the Effect on Quality of Electronic Health Records with Health Information Exchange (N = 18)

	Metric Description	Original Metric Set*
	Asthma	
1	The percentage of patients 18–56 years of age who were identified as having persistent asthma and who were appropriately prescribed medication during the measurement year†	NQF
	Cardiovascular Disease	
2	Percentage of patients hospitalized with AMI (acute myocardial infarction) who received persistent beta-blocker treatment (6 months after discharge)	AQA
3	Patients with ischemic vascular disease who have documentation of use of aspirin or another antithrombotic during the 12-month measurement period	NQF
4	Patients with ischemic vascular disease whose most recent LDL-C had a result of less than 100mg/dL	
	Congestive Heart Failure	
5	Percentage of patients with HF who also have paroxysmal or chronic atrial fibrillation who were prescribed warfarin therapy	DOQIT
	Diabetes	
6	Percentage of patients 18–75 years of age with diabetes whose most recent HbA1c level during the measurement year is > 9.0%	NQF
7	Percentage of patients 18–75 years of age with diabetes who had one or more HbA1c test(s) during the measurement year	NQF
8	Percentage of diabetic patients who had at least one HbA1C measured in the reporting period below 7%	TCNY
9	2-part measure: percentage of patients 18–75 years of age with diabetes whose most recent LDL-C level during the measurement year is < 130 mg/dL; percentage of patients 18–75 years of age with diabetes whose most recent LDL-C level during the measurement year is < 100 mg/dL	NQF
	Medication/Allergy Management	
10	Percentage of patients having documentation of current medication list in outpatient record	NQF
11	Percentage of patients having documentation of allergies and adverse reactions in patient record	NQF
	Mental Health	
12	Percentage of patients 18 years of age and older who had a follow-up visit within 30 days after being discharged for an inpatient mental health stay (including hospitalizations for depression, schizophrenia, attention deficit disorder, and personality disorders)‡	NCQA
	Osteoporosis	
13	Percentage of patients aged 50 years and older with fracture of the hip, spine, or distal radius who had a central dual-energy x-ray absorptiometry (DXA) measurement ordered or performed or pharmacologic therapy prescribed	PQRI
	Prevention	
14	The percentage of women 21–64 years of age who received one or more Pap tests to screen for cervical cancer	NQF
15	The percentage of women 40–69 years of age who had a mammogram to screen for breast cancer	NQF
16	The percentage of patients 65 years and older who ever received a pneumococcal vaccination	NQF
17	Flu shots for adults (50–64): the percentage of patients 50–64 years who received an influenza vaccination; flu shots for older adults: the percentage of patients 65 years and older who received an influenza vaccination	NQF
18	Colorectal cancer screening by colonoscopy performed (age 50–80)	TCNY

* NQF, National Quality Forum; AQA, Ambulatory Quality Alliance; LDL-C, low-density lipoprotein cholesterol; HF, heart failure; DOQIT, Doctor's Office Quality Information Technology; TCNY, Take Care New York; HbA1C, glycosolated hemoglobin; NCQA, National Committee on Quality Assurance; PQRI, Physician Quality Reporting Initiative. Please see Table 1 for references for each metric set.

† This metric was modified to include only adult patients.

‡ Because of some concerns about the way it was worded, the metric from the Center for Quality Assessment and Improvement in Mental Health (CQAIMH) set was substituted with the wording from a similar metric in the Healthcare Effectiveness and Data Information Set (HEDIS). The HEDIS version was modified to include only adult patients.

ing metrics: test ordering, medication management, referrals, follow-up after discharge, and revisits. We developed 14 metrics to address these topics (Table 3, pages 366–367).

The three metrics on test ordering addressed repeat blood tests, imaging studies, and cardiac studies, specifying different time intervals for “repeat” testing of different types. The metrics further specified that if a test had been done for a particular patient, the next test of the same type within the repeat interval counted as a repeat test, regardless of the ordering provider. Thus, the metrics took the perspective of the community or payer, rather than the individual physician, and required coordination of care across providers for the same patient.

The four metrics on medication management covered generic prescribing, formulary prescribing, availability of fill data, and ambulatory documentation of discharge medication lists from recent hospitalizations. All of these captured explicitly functionalities being employed by community-based HIT systems (that is, electronic alerts for generic prescribing and formulary adherence, provision of fill data, and electronic delivery of discharge medication lists).

The two metrics on referrals captured both sending information to a specialist and receiving the specialist’s recommendation. The two metrics on follow-up after discharge included follow-up visits and receipt of discharge summaries. The three metrics on revisits included emergency department visits and hospitalizations for ambulatory care-sensitive conditions, as well as re-admissions. Many of these concepts have been discussed by national quality organizations but have not yet appeared in metric sets.³³

LESSONS LEARNED FROM DEVELOPING METRICS DE NOVO

Many of the de novo metrics addressed the portion of quality that overlapped with utilization and that targeted coordination of care much more explicitly than existing metrics.

With respect to repeat testing, the literature was scarce on both quantifying appropriate frequencies for repeat testing and articulating exceptions. On the basis of existing literature, we were able to provide more specific test-frequency recommendations for blood tests than for imaging or cardiac testing.³⁴

Generic substitution and formulary compliance were hypothesized to be important drivers of financial savings expected from EHRs with health information exchange. Similarly, availability of fill data was hypothesized to augment efforts to promote adherence to medication, which would improve quality and potentially reduce cost. The metric on medication documentation was one step toward medication

reconciliation but did not require reporting of a physician’s act of reconciling different medication lists, which is being captured electronically in only a few large academic medical centers and not in community-based practices.³⁵

The metrics on referrals and follow-up after discharge reflected basic components of coordination of care that are rarely measured currently, given the challenges of capturing these in a paper-based system. Experts noted that EHRs and health information exchange introduce opportunities to measure communication between generalists and specialists in more systematic ways.

The metrics on revisits could potentially capture downstream effects of coordination of care. Because revisits are relatively rare, generating reliable estimates of them may require pooling the experiences of several physicians and measuring effects at the practice or community level.

Discussion

We developed a novel metric set that is potentially suitable for electronic reporting and is sensitive to the types of quality improvement expected from EHRs in the context of health information exchange. The total set of 32 metrics consists of 18 metrics developed and endorsed by national organizations plus 14 metrics developed de novo. Many of the de novo metrics directly address coordination of care.

Our metric set is distinct from the few other metric sets that evaluate specific aspects of coordination of care. The Care Transitions Measures elicit the patient’s perspective on coordination of care, which may be complementary to our metric set that elicits the physician’s perspective.²⁴ Some of the metrics included in Assessing Care of Vulnerable Elders (ACOVE) do target the physician’s perspective; however, many of them require that a physician “acknowledge” certain events, an action which is not easily amenable to electronic reporting.^{20,21}

Our metric set raises separate implications for measuring coordination of care and for electronic reporting of quality metrics. Our work suggests that the ability to measure coordination of care depends in part on the stakeholders and specific technologies involved in a community’s initiative. This is because stakeholders and technologies determine the types of data available for reporting. For example, communities that have pharmacy benefit manager data integrated into their health information exchanges will be able to report fill data and to measure adherence to medication, whereas other communities will not. Similarly, communities that implement structured medication reconciliation (a computer-assisted method for comparing multiple medication lists) will be able to measure

Table 3. Metrics Developed De Novo for Measuring with Electronic Reporting the Effect on Quality of Electronic Health Records with Health Information Exchange (N = 14)*

Test Ordering	
1	<p>Repeat Blood Tests For each type of blood test below, consider: Of all the tests ordered by a provider over a six-month period, how many represent tests for which results were already completed for that patient (regardless of the ordering provider) and are less than [insert appropriate repeat interval] old at the time of the second test? Hemoglobin (10 days) Creatinine (10 days) Sodium (10 days) Total cholesterol (6 weeks) HDL cholesterol (6 weeks) Thyroid stimulating hormone (6 weeks) Liver function tests (ALT/AST) (6 weeks) Ferritin (8 weeks) Hemoglobin A1c (12 weeks)</p>
2	<p>Repeat Imaging Studies Of those imaging studies (x-rays, ultrasounds, CT scans, and MRIs) ordered by a provider over a three-month period, how many represent tests for which results were already completed for that patient (regardless of the ordering provider) and are no more than 60 days old at the time of the second test?</p>
3	<p>Repeat Cardiac Studies Of those cardiac studies (all variants of stress tests and echocardiography) ordered by a provider over a three-month period, how many represent tests for which results were already completed for that patient (regardless of the ordering provider) and are no more than 90 days old at the time of the second test?</p>
Medications	
4	<p>Generic Prescribing Of all medications prescribed by a given provider and filled by patients over a three-month period, how many are filled as generic?</p>
5	<p>Formulary Prescribing Of all medications prescribed by a given provider and filled by patients over a three-month period, how many are on formulary?</p>
6	<p>Fill Data Of all patient visits to a given provider over a three-month period, how many have fill data available at the point of care?</p>
7	<p>Discharge Medication Documentation Of all patients discharged from a hospital over a three-month period, how many patients have the discharge medication list documented by the primary care provider during the first outpatient visit following discharge (which could occur up to one month after the discharge)?</p>
Referrals	
8	<p>Reason for Referral Of all patients referred to a specialist by a primary care physician over a three-month period, how many have the primary care physician's reason for the referral sent to the specialist's office?</p>
9	<p>Specialist Recommendations Of all patients referred to and seen by a specialist for a given primary care physician, for how many patients is the specialist's recommendations sent back to the primary care physician by the time of the patient's next follow-up appointment with the primary care physician (which could occur up to three months after the referral visit)?</p>
Follow-up After Discharge	
10	<p>Post-Discharge Hospital Follow-Up Of all patients who are hospitalized in a three-month period, how many are seen by their primary care physicians within 14 days of discharge?</p>
11	<p>Hospital Discharge Summary Of all patients who are hospitalized in a three-month period, how many had a discharge summary received by their primary care physician within 14 days of discharge?</p>

(continued on page 367)

Table 3. Metrics Developed De Novo for Measuring with Electronic Reporting the Effect on Quality of Electronic Health Records with Health Information Exchange (N = 14)* (continued)

	Revisits
12	Ambulatory Care–Sensitive Conditions ED Visits[†] Of all patients with at least one ambulatory care–sensitive condition that are in a physician’s panel at a given time, how many are seen in the emergency department over the subsequent three months for that ambulatory care–sensitive condition?
13	Ambulatory Care–Sensitive Conditions Hospitalizations Visits[†] Of all patients with at least one ambulatory care–sensitive condition that are in a physician’s panel at a given time, how many are hospitalized over the subsequent three months for that ambulatory care–sensitive condition?
14	Re-admissions Of all patients who are hospitalized in a three-month period, how many are re-admitted within 30 days of discharge?

* HDL, high-density lipoprotein; ALT, alanine aminotransferase; AST, aspartate aminotransferase; Hemoglobin A1c, glycosylated hemoglobin; CT, computerized tomography; MRI, magnetic resonance imaging; ED, emergency department.

[†] Common ambulatory care–sensitive conditions for adults, as defined by the Agency for Healthcare Research and Quality, include chronic obstructive pulmonary disease, adult asthma, congestive heart failure, angina without procedure, hypertension, diabetes with long-term complications, diabetes with short-term complications, lower extremity amputations among patients with diabetes, uncontrolled diabetes without complication, bacterial pneumonia, and urinary tract infections requiring hospitalization. (Source: Reference 33. Kruzikas D.T., et al.: *Preventable Hospitalizations: Window into Primary and Preventive Care 2000*. <http://www.ahrq.gov/data/hcup/factbk5/> (last accessed May 15, 2009).)

adherence to guidelines for medication reconciliation, but other communities will not.

The ability to measure coordination of care also depends on the degree of electronic integration of a community, because this affects the accuracy of reporting. For example, a well-integrated community will be able to report rates of certain tests more accurately than a less well-integrated community, because health information exchange would be required to integrate electronically results from tests performed only by specialists into the EHRs of generalists.

Our work raises three implications with respect to electronic quality reporting. First, capabilities for reporting are evolving and vary across EHR products. We aimed to propose a metric set that would be feasible for reporting now. However, we acknowledge that metrics that seem quite basic from a quality standpoint may still be remarkably challenging to report electronically from clinical data. Having a metric set like this one may, in fact, help “push the envelope” of what vendor systems can do.

Second, to report metrics electronically, one needs to “translate” existing specifications into those appropriate for clinical data sources. Specifications operationalize numerators, denominators, and exclusion criteria for each metric. Specifications for existing metrics were typically designed for claims data and relied on billing codes to identify the presence of disease. EHRs and health information exchanges offer additional data for identifying disease, including problem lists, abnormal laborato-

ry values, and medication lists. When these data are used, more patients are identified as eligible for inclusion in a metric than when claims are used alone.³⁶ Furthermore, the technical approach for converting clinical data into an electronic report is currently vendor specific, which may impede widespread, standardized electronic reporting because of the effort required by each vendor. National efforts to standardize EHR data with quality reporting in mind are under way, including work by the National Quality Forum’s Health Information Technology Expert Panel (HITEP)³⁷ and the Healthcare Information Technology Standards Panel (HITSP).³⁸

Third, the validity of electronic reporting may depend in part on physicians’ documentation, regardless of specifications. For example, the metric on aspirin use in patients with ischemic vascular disease depends on physician documentation, because aspirin is available over-the-counter. Over-the-counter medications are frequently missing from medication lists both because physicians are not required to prescribe them and because patients may not report them.³⁹

This study has several limitations. It is possible that we overlooked existing metric sets relevant to our research; however, our national experts felt that both the metric sets and the individual metrics were complete. Ideally, we would have selected metrics already proven to distinguish high- and low-performing systems; however, because such studies have generally not yet been conducted, we relied on expert opinion. We used two different sets of rating scales in Round One and Round Two,

which could potentially have affected the selection of the final metrics; however, Round Two essentially preserved the domains contained in Round One and added other, more detailed domains that were not feasible to include in Round One (because of the much higher number of metrics being considered in Round One). We acknowledge that the metric set consists mostly of process measures and a few intermediate clinical outcomes; if these are shown to be affected by EHRs with health information exchange, future metric sets could include long-term clinical outcomes. Our selection of de novo metrics was influenced by our collective experiences with multiple community-based initiatives, which may not reflect the experience of every community nationally, although we worked with national experts to make the metric set broad and widely relevant. The de novo metrics do not yet have detailed specifications for retrieval, and these need to be developed. Future studies should test and, if needed, improve the reliability of electronic retrieval of the metrics, compared to chart review and to claims. Once reliability of electronic retrieval is achieved, these metrics would require further validation through studies determining the effects of EHRs with health information exchange on metric performance.

Future generations of this metric set could shift quality measurement more completely into the realm of electronic clinical data sources, including metrics that could not be duplicated by claims or duplicated only with great effort through manual chart review. These would include metrics that (1) capture appropriateness of medical decision making, such as appropriate escalation of the intensity of medications for patients with chronic disease⁴⁰; (2) incorporate patient preferences by excluding those who refuse certain interventions; and (3) consider patients' comorbidities and factor in physicians' and patients' prioritization of these comorbidities when assessing adherence to guidelines.⁴¹ Some of these future metrics may require modifications in software, such that structured rather than free-text fields are created for the variables needed for reporting. By articulating the direction of quality measurement, work of this kind may influence software design and configuration, EHR implementation and use, and methods of EHR data extraction.

Summary

We created a novel metric set that could be used to measure the potential quality effects expected from electronic integration of physician practices, hospitals, laboratories, pharmacies, and other sources of clinical data. The metric set is potentially suitable for electronic reporting and contains 18 metrics previous-

ly developed by national organizations and 14 newly developed. This metric set may have broad utility, given that EHRs with health information exchange are being disseminated widely. This work also has the potential to stimulate discussion about how to improve the way data are entered and extracted from clinically rich electronic data sources, with the ultimate goal of more accurately measuring and improving the quality of ambulatory care today. ■

This work was supported by the Agency for Healthcare Research and Quality (AHRQ grant #R18 HS 017067). The authors thank the members of the national expert panel for their participation. The authors also thank David Blumenthal, M.D., M.P.P., for co-moderating the national expert panel meeting with Harold Pincus, M.D. Earlier versions of this work were presented at the AHRQ Annual Meeting on September 8, 2008, in Bethesda, Maryland, and at the National Meeting of the American Medical Informatics Association on November 11, 2008, in Washington, DC.

Lisa Kern, M.D., M.P.H., is Assistant Professor of Public Health and Medicine, Department of Public Health, Weill Cornell Medical College (WCMC), New York City; Deputy Director, Health Information Technology Evaluation Collaborative (HITEC), New York City. **Rina Dhopeshwarkar, M.P.H.**, is Research Coordinator, Department of Public Health, WCMC. **Yolanda Barrón, M.S.**, is Research Associate in Biostatistics, Department of Public Health, WCMC. **Adam Wilcox, Ph.D.**, is Assistant Professor of Biomedical Informatics, Columbia University, New York City; Director of Clinical Databases, NewYork-Presbyterian Hospital (NYPH), New York City. **Harold Pincus, M.D.**, is Professor and Vice Chair for Strategic Initiatives, Department of Psychiatry, Columbia University; Director of Quality and Outcomes Research, NYPH; Senior Scientist, RAND Corporation; and Associate Director, Irving Institute for Clinical and Translational Research, Columbia University. **Rainu Kaushal, M.D., M.P.H.**, is Associate Professor of Pediatrics, Medicine and Public Health, WCMC; Chief of the Division of Quality and Medical Informatics, Department of Pediatrics; Director of Pediatric Quality and Safety at Komansky Center for Children's Health, NYPH; Executive Director, HITEC; and a member of *The Joint Commission Journal on Quality and Patient Safety's* Editorial Advisory Board. Please address requests for reprints to Rainu Kaushal, M.D., M.P.H., rak2007@med.cornell.edu.

Online-Only Content

See the online version of this article for Appendix 1. National Expert Panel Members.

References

1. Smith P.C., et al.: Missing clinical information during primary care visits. *JAMA* 293:565–571, Feb. 2, 2005.
2. Forster A.J., et al.: The incidence and severity of adverse events affecting patients after discharge from the hospital. *Ann Intern Med* 138:161–167, Feb. 4, 2003.
3. Institute of Medicine: *To Err Is Human: Building a Safer Health System*. Washington, DC: National Academy Press, 2000.
4. Agency for Healthcare Research and Quality: *Closing the Quality Gap: A Critical Analysis of Quality Improvement Strategies: Volume 7—Care Coordination, Structured Abstract*, Jun. 2007. <http://www.ahrq.gov/clinic/tp/caregapt.htm> (last accessed May 15, 2009).
5. Institute of Medicine: *Priority Areas for National Action: Transforming Health Care Quality*. Washington, DC: National Academies Press, 2003.
6. Pham H.H., et al.: Primary care physicians' links to other physicians through Medicare patients: The scope of care coordination. *Ann Intern Med* 150:236–242, Feb. 17, 2009.
7. American Recovery and Reinvestment Act of 2009, Pub L No. 111-5, 123 Stat 115.
8. Foundation of Research and Education of American Health Information Management Association: *State-Level Health Information Exchange. Final Report I: Roles in Ensuring Governance and Advancing Interoperability*. Mar. 10, 2008. <http://www.shie.org/Docs/FinalReportPart1.8.pdf> (last accessed May 15, 2009).
9. Blumenthal D., Glaser J.P.: Information technology comes to medicine. *N Engl J Med* 356:2527–2534, Jun. 14, 2007.
10. Healthcare Information Management and Systems Society: *HIMSS Electronic Health Record Definitional Model, Version 1.1*, Sep. 24, 2003. <http://www.himss.org/content/files/ehrattributes070703.pdf> (last accessed May 15, 2009).
11. National Alliance for Health Information Technology: *NAHIT Releases HIT Definitions*, http://www.nahit.org/pandc/press/pr5_20_2008_1_33_49.asp (last accessed May 15, 2009).
12. Institute of Medicine: *Performance Measurement: Accelerating Improvement*. Washington, DC: National Academies Press, 2006.
13. New York State Department of Health: *Health Information Technology (Health IT) Grants—HEAL NY Phase I*, May 2006. <http://www.health.state.ny.us/technology/awards/> (last accessed May 15, 2009).
14. New York State Department of Health: *New York Awards \$105 Million in Health Information Technology Grants*. Mar. 28, 2008. http://www.ny.gov/governor/press/press_0328081.html (last accessed May 13, 2009).
15. Kern L.M., et al.: Promoting interoperable health information technology in New York State. *Health Aff (Millwood)* 28:493–504, Mar.–Apr. 2009.
16. AQA Alliance: *Starter Set of 26 Measures Endorsed by AQA*, Apr. 2006. <http://www.aqaalliance.org/performancewg.htm> (last accessed May 15, 2009).
17. National Committee on Quality Assurance: *Healthcare Effectiveness Data and Information Set*, 2005. <http://www.ncqa.org/Portals/0/HEDISQM/Archives/2005/MeasuresList.pdf> (last accessed May 15, 2009).
18. National Quality Forum: *NQF-Endorsed National Voluntary Consensus Standards for Physician-Focused Ambulatory Care*, 2006. http://www.qualityforum.org/pdf/reports/ambulatory_care.pdf (last accessed May 15, 2009).
19. New York City Department of Health and Mental Hygiene: *Take Care New York*, 2004. <http://www.nyc.gov/html/doh/html/tcny/index.shtml> (last accessed May 15, 2009).
20. RAND: *ACOVE 2 Quality Indicators*, 2005. <http://www.rand.org/health/projects/acove/acove2/> (last accessed May 15, 2009).
21. Coleman E.A.: Appendix I: Transitional care performance measurement. In: Institute of Medicine: *Performance Measurement: Accelerating Improvement*. Washington, DC: National Academies Press, 2006, pp. 250–286.
22. Centers for Medicare & Medicaid Services: *Doctor's Office Quality—Information Technology (DOQ-IT): Ambulatory care measures*. <http://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier2&cid=1143577171055> (last accessed May 15, 2009).
23. Centers for Medicare & Medicaid Services: *2007 Physician Quality Reporting Initiative (PQRI) Physician Quality Measures*, 2007. <http://www.cms.hhs.gov/PQRI/Downloads/2007PQRIMeasuresList.pdf?agree=yes&next=Accept> (last accessed May 15, 2009).
24. Coleman E.A.: *The Care Transitions ProgramSM*. <http://www.caretransitions.org/downloadctm.asp> (last accessed May 15, 2009).
25. National Committee on Quality Assurance: *Standards and Guidelines for Physician Practice Connections—Patient-Centered Medical Home (PPC-PCMH)*, 2008. http://www.cms.hhs.gov/DemoProjectsEvalRpts/downloads/MedHome_PPC.pdf (last accessed May 15, 2009).
26. Bonow R.O. et al: ACC/AHA clinical performance measures for adults with chronic heart failure: A report of the American College of Cardiology/American Heart Association Task Force on Performance Measures (Writing Committee to Develop Heart Failure Clinical Performance Measures) endorsed by the Heart Failure Society of America. *J Am Coll Cardiol* 46:1144–1178, Sep. 20, 2005.
27. American Medical Association: *Clinical Performance Measures – Chronic Stable Coronary Artery Disease: Tools Developed by Physicians for Physicians*, 2005. <http://www.ama-assn.org/ama/no-index/physician-resources/15392.shtml> (last accessed May 15, 2009).
28. National Committee on Quality Assurance: *Diabetes Physician Recognition Program*, 2008. <http://www.ncqa.org/tabid/139/Default.aspx> (last accessed May 15, 2009).
29. American Geriatrics Society: *American Geriatrics Society Pay for Performance Proposal*, 2006. http://www.americangeriatrics.org/policy/2006p4p_proposalPF.shtml (last accessed May 15, 2009).
30. Center for Quality Assessment and Improvement in Mental Health: *National Inventory of Mental Health Quality Measures*, 2005. <http://www.cqaimh.org/NIMHQM.htm> (last accessed May 15, 2009).
31. Center for Quality Assessment and Improvement in Mental Health: *The STAndards for Bipolar Excellence Project*, 2006. <http://www.cqaimh.org/stable.html> (last accessed May 15, 2009).
32. McGlynn E.A. et al.: *The Quality of Health Care Delivered to Adults in the United States: Technical Appendix, RAND Health Working Paper*, 2006. http://www.rand.org/pubs/working_papers/WR174-1/ (last accessed May 15, 2009).
33. Kruzikas D.T., et al.: *Preventable Hospitalizations: A Window into Primary and Preventive Care*, 2000. <http://www.ahrq.gov/data/hcup/factbk5/> (last accessed May 15, 2009).
34. van Walraven C., Raymond M.: Population-based study of repeat laboratory testing. *Clin Chem* 49:1997–2005, Dec. 2003.
35. Poon E.G., et al.: Design and implementation of an application and associated services to support interdisciplinary medication reconciliation efforts at an integrated healthcare delivery network. *J Am Med Inform Assoc* 13:581–592, Nov.–Dec. 2006.
36. Tang P.C., 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* 14:10–15, Jan.–Feb. 2007, Epub Oct. 26, 2006.
37. National Quality Forum: *Health Information Technology Expert Panel II*, 2009. <http://www.qualityforum.org/projects/ongoing/HITEP/> (last accessed May 15, 2009).
38. American National Standards Institute: *Healthcare Information Technology Standards Panel*, 2008. <http://www.hitsp.org> (last accessed May 15, 2009).
39. Nassaralla C.L., et al.: Implementation of a medication reconciliation process in an ambulatory internal medicine clinic. *Qual Saf Health Care* 16:90–94, Apr. 2007.
40. Kerr E.A., et al.: Avoiding pitfalls in chronic disease quality measurement: A case for the next generation of technical quality measures. *Am J Manag Care* 7:1033–1043, Nov. 2001.
41. Boyd C.M., et al.: Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: Implications for pay for performance. *JAMA* 294:716–724, Aug. 10, 2005.

Appendix 1. National Expert Panel Members

These national expert panel members validated metrics for the potential quality effects of ambulatory electronic health records with health information exchange.

Kristine Martin Anderson, M.B.A., Principal, Booz Allen Hamilton, Washington, DC

Steven Asch, M.D., M.P.H., Deputy Associate Chief of Staff, Health Services Research and Development, Department of Veterans Affairs, Greater Los Angeles Healthcare System, Los Angeles; Associate Professor of Medicine, David Geffen School of Medicine at UCLA, Los Angeles; and Consultant, RAND Corporation, Los Angeles

Randolph Barrows, M.S., M.D., formerly Chief Medical Information Officer, MedAllies, Fishkill, New York

David Bates, M.D., M.Sc., Chief, Division of General Internal Medicine, Brigham and Women's Hospital, Boston; Medical Director of Clinical and Quality Analysis, Partners Healthcare, Boston; Professor of Medicine, Harvard Medical School, Boston; and Professor of Health Policy and Management, Harvard School of Public Health, Boston

John Blair, III, M.D., President, Taconic IPA, Fishkill, New York; and CEO, MedAllies, Fishkill

Rachel Block, formerly Executive Director, New York eHealth Collaborative, New York City; currently Deputy Commissioner, Office of Health Information Technology Transformation, New York State Department of Health, Albany, New York

David Blumenthal, M.D., M.P.P., formerly Director of Institute for Health Policy at Massachusetts General Hospital, Boston; and Samuel O. Thier Professor of Medicine and Professor of Health Care Policy, Harvard Medical School, Boston; currently National Coordinator for Health Information Technology, Department of Health and Human Services, Washington, DC

Helen Burstin, M.D., M.P.H., Senior Vice President for Performance Measures, The National Quality Forum, Washington, DC

David Cohen, M.D., Vice Chair, Medicine & Sr VP, Clinical Integration & Affiliations, Maimonides Medical Center, Brooklyn, New York

Jill Eisenstein, M.S., Manager of Business Operations, Greater Rochester RHIO, Rochester, New York

Lori Evans, M.P.P., M.P.H., formerly Deputy Commissioner, Office of Health Information Technology Transformation, New York State Department of Health, Albany, New York

Marybeth Farquhar, M.S.N., R.N., Agency for Healthcare Research and Quality, Rockville, Maryland

Ellen Flink, M.B.A., Director of Research in Patient Safety and Quality Initiatives, New York State Department of Health, Albany, New York

Mark Frisse, M.D., M.B.A., M.Sc., Accenture Professor of Biomedical Informatics; Director, Regional Informatics Programs, Vanderbilt Center for Better Health, Nashville, Tennessee

Martin Hickey, M.D., formerly Sr. VP Health Care Affairs, Excellus BlueCross BlueShield, Rochester, New York; currently President & CEO, Alegent Health Clinic, Omaha

George Hripcsak, M.D., M.S., Professor of Biomedical Informatics and Chair, Department of Biomedical Informatics, Columbia University, New York City

Rainu Kaushal, M.D., M.P.H., Chief, Division of Quality and Medical Informatics, Department of Pediatrics, Weill Cornell Medical College, New York City; Executive Director, Health Information Technology Evaluation Collaborative, New York City; and Director of Quality, Komansky Center for Children's Health, NewYork-Presbyterian Hospital, New York City

Karen Kmetik, Ph.D., Director, Clinical Performance Evaluation, American Medical Association, Chicago

Irene Koch, Esq., Executive Director, Brooklyn Health Information Exchange, Brooklyn, New York

Ted Kremer, M.P.H., Executive Director, Greater Rochester RHIO, Rochester, New York

Alex Low, Health Policy Analyst, United Hospital Fund/New York eHealth Collaborative, New York City

John Morley, M.D., F.A.C.P., Medical Director, Office of Health Systems Management, New York State Department of Health, Albany, New York

Alvin Mushlin, M.D., Sc.M., Chairman of Department of Public Health, and Nanette Laitman Distinguished Professor of Public Health, Weill Cornell Medical College, New York City; and Public Health Physician-in-Chief, NewYork-Presbyterian Hospital, New York City

J. Marc Overhage, M.D., Ph.D., Director of Medical Informatics, Regenstrief Institute; and Regenstrief Professor of Medicine, Indiana University School of Medicine, Indianapolis, Indiana

Greg Pawlson, M.D., M.P.H., Executive Vice President, National Committee for Quality Assurance, Washington, DC

Daniel Porreca, Executive Director, HealtheLink, Buffalo, New York

(continued on next page)

Online-Only Content 

Appendix 1. National Expert Panel Members (continued)

Harold Pincus, M.D., Professor and Vice Chair for Strategic Initiatives, Department of Psychiatry, College of Physicians and Surgeons, Columbia University, New York City; Associate Director, Irving Institute for Clinical and Translational Research, Columbia University; Director of Quality and Outcomes Research, New York-Presbyterian Hospital, New York City; and Senior Scientist, RAND Corporation, Santa Monica, California

Neil Powe, M.D., M.P.H., M.B.A., formerly Professor of Medicine and Epidemiology, Johns Hopkins University, Baltimore; Director, Welch Center for Prevention, Epidemiology and Clinical Research, Baltimore; Director, Johns Hopkins Clinical Research Scholars Program, Baltimore; and Director, Johns Hopkins Pre-doctoral Clinical Research Training Program; currently Professor and Vice Chair of Medicine, University of California at San Francisco, San Francisco; and Chief of Medical Services, San Francisco General Hospital, San Francisco

Raghu Ram, M.D., Chair, Clinical Advisory Committee, HealthLink, Buffalo, New York

Haya Rubin, M.D., Ph.D., Adjunct Professor, Johns Hopkins University, Baltimore

Bill Schroth, M.B.A., Consultant, New York State Department of Health, Albany, New York

Anthony Shih, M.D., M.P.H., formerly Assistant Vice President, Program Director, Quality Improvement and Efficiency, The Commonwealth Fund, New York City; currently Chief Quality Officer and Vice President of Strategic Planning, IRPO, Lake Success, New York

Paul Tang, M.D., M.S., Chief Medical Information Officer, Palo Alto Medical Foundation, Palo Alto, California

Sheila Teasdale, M.Med.Sci., F.B.C.S., Clinical Performance Evaluation, American Medical Association, Chicago

Micky Tripathi, Ph.D., M.P.P., President and CEO, Massachusetts eHealth Collaborative, Waltham, Massachusetts

Paul Wallace, M.D., Medical Director for Health and Productivity Management Programs; and Senior Advisor for the Care Management Institute and Avivia, Kaiser Permanente, Oakland, California