How is the electronic health record being used? Use of EHR data to assess physician-level variability in technology use

Jessica S Ancker,[1,2] Lisa M Kern,[1,2] Alison Edwards,[1,2] Sarah Nosal,[3] Daniel M Stein,[1] Diane Hauser,[3] Rainu Kaushal,[1,2] with the HITEC Investigators

▶ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/amiajnl-2013-002627)

[1]Department of Healthcare Policy and Research, Center for Healthcare Informatics and Policy, Weill Cornell Medical College, New York, USA
[2]Health Information Technology Evaluation Collaborative (HITEC), New York, USA
[3]Institute for Family Health, New York, USA

Correspondence to
Dr Jessica Ancker, 425 E. 61st Street, Suite 301, New York, NY 10065, USA; jsa7002@med.cornell.edu

Received 31 December 2013
Revised 15 April 2014
Accepted 21 May 2014
Published Online First 9 June 2014

ABSTRACT

Background Studies of the effects of electronic health records (EHRs) have had mixed findings, which may be attributable to unmeasured confounders such as individual variability in use of EHR features.

Objective To capture physician-level variations in use of EHR features, associations with other predictors, and usage intensity over time.

Methods Retrospective cohort study of primary care providers eligible for meaningful use at a network of federally qualified health centers, using commercial EHR data from January 2010 through June 2013, a period during which the organization was preparing for and in the early stages of meaningful use.

Results Data were analyzed for 112 physicians and nurse practitioners, consisting of 430 803 encounters with 99 649 patients. EHR usage metrics were developed to capture how providers accessed and added to patient data (eg, problem list updates), used clinical decision support (eg, responses to alerts), communicated (eg, printing after-visit summaries), and used panel management options (eg, viewed panel reports). Provider-level variability was high: for example, the annual average proportion of encounters with problem lists updated ranged from 5% to 60% per provider. Some metrics were associated with provider, patient, or encounter characteristics. For example, problem list updates were more likely for new patients than established ones, and alert acceptance was negatively correlated with alert frequency.

Conclusions Providers using the same EHR developed personalized patterns of use of EHR features. We conclude that physician-level usage of EHR features may be a valuable additional predictor in research on the effects of EHRs on healthcare quality and costs.

INTRODUCTION

The long-term goal of the federal electronic health record (EHR) incentive program is to improve the quality and safety of healthcare.[1] To date, however, research on the effects of EHRs on healthcare delivery has been highly mixed. While many studies have found associations between EHRs and quality improvement, others have found adverse effects, and still others have shown no effect.[2–9] We[10] and others[8] have argued that the actual effects of EHRs in these studies may have been obscured by a methodological issue, that is, measuring EHR presence or absence as a binary factor. EHRs are complex pieces of technology offering multiple functions and features, and they are embedded in and shaped by complex social environments.[11] The same EHR product may be customized differently in different organizations, and implementation processes and organization-specific workflows can also affect how certain features are used. A simple example is that electronic order sets are often developed or modified by healthcare organizations on the basis of local clinical priorities, and thus can vary at the level of the practice or department. Furthermore, it is also highly likely that individual physicians vary in their use of EHR features as a result of preferences or experience. For example, some physicians may habitually use a particular order set as a way of simplifying the ordering process, whereas others may be unaware of it or avoid using it because they disagree with its content or find its design unusable.

For these reasons, effects of EHRs may depend in part on the way EHRs are used by individual clinicians, not merely on whether EHRs are available. One way of assessing individual-level use of EHRs is the meaningful use (MU) 'objective measures,' which are intended to capture and promote certain types of EHR use (eg, recording demographics). Another approach is self-report, as has been used in surveys[8 9 12 13] in which physicians characterize their use of EHR features such as the problem list or radiology result delivery. Lanham and colleagues recently employed interviews and direct observation to distinguish between intensive and less intensive EHR use.[14] However, the availability of EHR data itself creates possibilities for objectively measuring EHR use, capturing granular metrics of usage that could be scaled up or even automated. For example, data capture directly from clinical decision support (CDS) systems has frequently been analyzed to assess rates of response to alerts and reasons for overrides.[15–17] As a number of other commentators have noted, diffusion of innovation (adoption of the technology) differs from infusion (the intensity or sophistication with which the technology is used).[18]

Our objective in the current study was to analyze data from an EHR to capture individual physician use of a wider array of EHR functions. We hypothesized that the EHR data would reveal physician-level variation in use of EHR functions, that some of the variation would be linked to provider characteristics, but also that intensity of use of EHR functions would increase over a time period coinciding with the organization's preparation for and early engagement in MU. Finding a high degree of physician-level variability would suggest a high degree of personalization that would

To cite: Ancker JS, Kern LM, Edwards A, *et al*. *J Am Med Inform Assoc* 2014;**21**:1001–1008.

support the potential value of measuring physician-level usage patterns in studies of EHR effects. Conversely, lack of variability might suggest that individual usage patterns would not add information to studies of EHR effects.

METHODS

Setting

The Institute for Family Health (IFH) is a network of federally qualified health centers (FQHCs) providing safety net care in New York City and less urbanized regions to the north of the city. IFH offers primary care services at 18 sites, with more than 100 physicians (almost all family practitioners) and a patient population of approximately 100 000. IFH has been using the EpicCare EHR since 2003 and is currently a level III patient-centered medical home. EHR customization has taken place at the institutional level and does not vary by clinic, but in some cases clinics may have staffing differences that affect EHR workflow. About half of IFH's family practitioners attested to MU stage I under the Medicaid program in 2012. During the period of the study, two new clinics joined the network, resulting in an increase in numbers of providers and patients.

Metric development

A team of researchers with diverse training (informatics, medicine, biostatistics, and health services research) conducted a literature review and developed a conceptual model broadly categorizing mechanisms through which the EHR might be expected to influence individual provider behavior to affect healthcare quality in the ambulatory setting. The research team then worked with an EHR database analyst at IFH to extract data and construct sample metrics within each of the categories. In this study, we report on 19 sample metrics in five categories (table 2):

1. Use of functions pertaining to the completeness of or access to patient data[1 8 19–23] (eg, frequency with which providers updated problem lists; time to access test results)
2. Use of decision support[1 19 24–26] (eg, responses to pushed alerts, reminders, and best practices; also included use of pulled or passive decision support in the form of order sets)
3. Use of electronic ordering (eg, laboratory tests and prescriptions)[9 27]
4. Use of features for care coordination and communication[1 19] (eg, secure messaging)
5. Use of panel-level reports[1 19 25 28] (tools to allow providers to see summarized data about an entire panel).

In the EHR system being studied, best-practice alerts were pushed to the provider at entry to the patient record with information about preventive care or other reminders. Prescribing alerts were interruptive and were pushed at order entry. Best-practice and prescribing alerts captured physician responses in different ways. For best-practice alerts, we considered the alert to be accepted if the provider clicked 'accept,' or opened the order set highlighted in the alert, or clicked to send a secure message through the EHR. For prescribing alerts, we considered the alert accepted if the provider indicated the drug had been discontinued or the order had been removed. (Alerts that were dismissed or overridden were considered overrides for the purpose of the current study, even if the provider gave a reason for the override.) We also classified best-practice alerts into disease/condition categories (asthma-related alerts, hyperlipidemia-related alerts, etc.). Prescribing alerts were classified by the vendor into types (drug–drug, drug–food, drug–allergy, drug–condition, and drug–alcohol) but not by pharmaceutical class or severity. Drug–alcohol alerts were not included in this study as these alerts were generally intended to be informative rather than to change prescribing behavior. The panel-level reporting function was implemented starting in 2011.

Study sample and data sources

The study sample included all IFH family practitioners who met the criteria for 'eligible provider' under the Medicaid MU program and who had at least one documented patient encounter between January 1, 2010 and June 30, 2013, as well as all patients seen by these providers. Each provider was included in the data set for the year during which he/she saw patients and therefore the included set of providers varied slightly year to year. We included both providers who attested to MU stage I in 2012 as well as those who had not done so by the time data collection was concluded. Encounter-level data for the same period were retrieved from the EHR database, including billing diagnoses, problem lists, orders, etc, as well as patient use of the electronic patient portal and messaging. Proprietary logic from the EHR vendor (following the criteria established by the Centers for Medicare & Medicaid Services (CMS)[29]) was used to create reports of the MU 'objective measures' for 2011, 2012, and 2013 for all included providers (the algorithms could not be applied to the 2010 data). In the current study, MU objective measures were reported for both the providers who attested and those who did not. To assess patient illness burden, the Johns Hopkins Aggregated Diagnosis Groups (ADG) count of comorbidities was computed.[30] Although the ADG algorithm was designed to be applied to claims data, we instead applied it to the encounter/billing diagnoses for each patient captured in the EHR (without considering any information in the EHR problem list).

The analysis of data from all clinicians allows conclusions about EHR use at the institutional level. To further explore individual-level changes over time, we repeated analyses on the subset of providers who contributed data to all 3.5 years of data collection.

Analysis

All metrics were reported at the provider level on the basis of all patient encounters during each calendar year from 2010 through June 2013 and also pooled across the study period. Categorical variables (such as frequency of updates to the problem list) were calculated as the proportion of encounters per provider. Continuous variables were reported as the mean and SD (for normally distributed variables) or median and IQR (for skewed variables) across encounters for that provider. For table 2, medians were reported for all variables because most were skewed. Descriptive statistics were used to examine the first hypothesis, that EHR data would reveal physician-level variation in use of EHR functions.

For the second set of hypotheses (that some of the variation would be linked to characteristics of the provider), we elected to reduce potential type I errors from multiple hypothesis testing by selecting a limited number of individual hypotheses to test. After discussion among the research team, we formulated three sets of hypotheses that appeared to have high face validity:

A. Response rates to decision-support and best-practice alerts would fall as alert receipt frequency rose, in a dose–response relationship. We hypothesized that providers who saw a particular condition or patient population frequently would be expected to become familiar with the standard of care and rely less upon support offered by the EHR. Conversely, providers who saw a particular condition less often might rely upon EHR guidance more frequently.

B. Alert response rates would be negatively correlated with panel size. Specifically, we hypothesized that information overload or alert fatigue would be likely to be associated with larger panel sizes and would likely reduce response rates.
C. Problem lists would be more likely to be updated for certain types of patients. Specifically, providers with sicker patients on average would be more likely to update problem lists as patients would have more medical problems. Providers with higher proportions of patients using the patient portal would be more likely to update problem lists because patients could use the portal to see their own problem lists. Finally, we hypothesized that providers would be more likely to update problem lists for patients they were seeing for the first time than for established patients as they would be more likely to perceive a need to update the patient record for new patients.

Hypotheses A, B, and C were assessed by generalized linear mixed models (GLMM) using a gamma distribution. Bivariate GLMM were adjusted for within-provider clustering and were weighted by year (ie, 2013 data were weighted at 0.5 because only 6 months of data were included). Multivariable models were additionally adjusted for MU attestation (yes vs no), provider gender, degree (Doctor of Medicine (MD)/Doctor of Osteopathic Medicine (PO) vs Advanced Practice Nurse (APN)), number of patients, and average ADG of the patient panel. Results are reported as incidence rate ratios (IRR) per 100 alerts or per 100 patients, interpreted as the extent to which the likelihood of the outcome changed for every 100 additional alerts or 100 additional patients.

For the third hypothesis, that intensity of use of EHR functions would increase over time, we performed tests for trend using GLMM, adjusted for clustering within physician. Data were modeled using gamma distributions for positively skewed data, beta distributions for negatively skewed data, or normal distributions. As data were available only for half of 2013, models were either weighted by year or else 2013 data were excluded for metrics that represented averages over a 12-month span (as indicated in table 2).

All analyses were conducted in SAS V.9.3 (Cary, North Carolina, USA).

Approvals

The study was approved by the institutional review boards of both Weill Cornell Medical College and the IFH. A waiver of consent was granted as providers and patients were both deidentified in the data sets.

RESULTS

The complete data set contained 112 family practitioners and 430 803 encounters with 99 649 unique patients (table 1). Forty-two providers had attested for MU by the completion of data collection, with the remaining providers planning to attest at future dates.

Metric variability

Median usage varied widely from metric to metric (table 2). For example, prescribing alerts were generally overridden, whereas best-practice alerts were accepted about 20% of the time. Problem lists and medication lists were rarely marked as reviewed (<10% for both), whereas allergies were marked as reviewed more than 75% of the time. Only about two patients were messaged per month, but patient-initiated secure messages were almost always responded to (median response rate per provider: 100%). Providers updated patient problem lists in a median of 20–25% of encounters, virtually always by adding problems rather than dropping them. The panel-level view, which allowed providers to build reports on the basis of their entire patient panel, was used by only a few individuals.

Also, as hypothesized, provider-level variability was also quite high in most of the metrics (table 2). Table 2 shows the IQR to represent the central half of the data, but the full ranges were much wider. For example, best-practice alert acceptance rates per provider ranged from 0% to 68%, the annual average proportion of encounters with the problem list updated ranged from 4.9% to 60.2% per provider, and the annual average proportion of drug–allergy alerts that prompted a drug discontinuation ranged from 0.0% to 62.5% per provider.

Associations with alert firing frequency and characteristics of providers and provider–patient relationship

Overall, as a provider received more best-practice alerts, his or her alert acceptance rate decreased modestly but statistically significantly. The multivariable adjusted IRR was 0.99 per 100 alerts ($p<0.001$), indicating that for every additional 100 alerts received, the provider accepted 1% fewer of them. Thus, a provider who received 7000 alerts (and about one quarter of providers received 7000 alerts or more annually) was 30% less likely to accept these alerts than a provider who received 4000 alerts, the approximate median for the group (table 3). Similar correlations held within 14 of the 18 categories of alerts evaluated in bivariate GLMM models; 13 of these associations remained significant in models that also adjusted for MU attestation, provider gender, degree, panel size, and average ADG of patient panel (table 3). Drug alerts showed a negative association between frequency and alert acceptance in the bivariate models, but the association was not statistically significant in the multivariable model . Response rates for best-practice alerts were not significantly correlated with panel size ($p>0.05$), but drug alerts were slightly less likely to result in drug discontinuation as panel size increased (multivariable adjusted IRR 0.94 per 100 patients; $p=0.005$).

Providers were more likely to update problem lists for a patient they were seeing for the first time than for a patient they had seen before (problem list updated for a median of 26.1% of new patient encounters compared to 20.5% of established patient encounters; $p<0.001$). Problem list updates were not significantly associated with proportion of patients using their portal account or with average ADG count for the patient panel.

Forty-two providers attested for MU in 2012, but only 41 saw patients at IFH in 2013 and were included in the 2013 data set. As shown in table 2, providers who attested for MU had higher scores than non-MU providers on four of the MU measures and lower scores on one. Almost all the remaining MU measures had ceiling effects, with median compliance rates higher than 95%. The MU providers also had higher usage than non-MU providers for a few of the novel metrics.

Change in usage metrics over time

Tests for trend including data from all providers showed that median usage on seven of the 19 novel metrics changed significantly over the study period (table 2). For example, the proportion of encounters with problem lists updated increased from 19.5% to 25.0% ($p=0.002$) and the median number of days required to complete a review of a new laboratory test result dropped from 4.9 to 4.0 ($p=0.01$). The proportion of patients using the portal decreased slightly in 2010–2012 (table 2),

probably due to portal enrollment not keeping pace with the increasing patient population as several new centers were added to the FQHC network. However, among portal users, there was a significant increase over time in the proportion who used their account frequently (defined as 15 times per year, which was the top quartile of use).

In addition, median performance on eight of the 14 MU objective measures also rose significantly. (Remaining measures that did not rise significantly either had ceiling effects or missing data. In the case of e-prescribing, the rate already exceeded the minimum established by MU at baseline.)

As the set of providers included in each year's data was somewhat different, these findings should be interpreted as institution-level trends not provider-level trends. Providers contributed a mean of 2.7 years (SD 1.3) of data to the data set. However, 46 providers contributed data to all 3.5 years of the study and were included in a longitudinal subset analysis (see online supplementary appendix). Trends of increasing EHR usage within this longitudinal subset were almost identical to trends for the institutional-level analysis with all providers, in that all the same trends were statistically significant with effect sizes of similar magnitude with the following exceptions: (1) days to review laboratory test results continued to show a reduction over time, but the effect was no longer statistically significant ($p=0.08$); and (2) the rate of problem list review, which did not change significantly over time for the larger group, significantly declined over time in the subset ($p=0.02$). In the subset analysis, two of the metrics could not be modeled due to insufficient sample size or lack of convergence of the model (proportion of patients using portal ≥15 times per year, and the MU medication list metric).

DISCUSSION

Primary care providers using the same EHR system in a network of clinics developed personalized approaches to their use of EHR features, which varied along dimensions ranging from frequency with which they updated patient problem lists to responsiveness to CDS alerts. Some of the variability was associated with provider or patient–provider characteristics; for example, providers behaved differently with new patients than with established patients, and responded differently to uncommon alert types than to common ones. A certain amount of individual-level variability was evident in the MU objective metrics, even

Table 1 Provider and patient characteristics

Characteristic	Total or weighted average across study period
Providers, n	112
Attested in 2012, n (%)	42 (37.5)
Encounters per provider, median (25th–75th percentile)	2867 (934–6148)
Unique patients seen per provider, median (25th–75th percentile)	1375 (605–2548)
ADG comorbidity count of panel, median (25th–75th percentile)	0.3 (0.2–0.4)
Female sex, n (%)	71 (63.4)
Credential, n (%)	
MD/DO	93 (83.0)
Advanced Practice Nurse	19 (17.0)
Patients, n	99 649
Encounters per patient, median (25th–75th percentile)	2 (1–5)
Mean age (SD)*	36 (20)
Insurance type, n (%)†	
Medicaid fee-for-service or managed care	33 339 (36.5)
Medicare fee-for-service or managed care	9635 (9.7)
Commercial	30 134 (30.2)
Uninsured/self-pay	19 394 (19.5)
Other	3905 (3.9)
Female sex, n (%)	58 694 (58.9)
Race, n (%)	
White	34 035 (34.2)
Black or African American	24 111 (24.2)
Other	27 695 (27.8)
Unreported/refused	13 808 (13.9)
Ethnicity, n (%)	
Hispanic or Latino	30 507 (30.6)
Not Hispanic or Latino	52 997 (53.2)
Not collected/unknown	16 145 (16.2)
Preferred language, n (%)	
English	79 369 (79.7)
Spanish	7757 (7.8)
Other language	475 (0.5)
Not collected/unknown/declined	12 048 (12.1)

Percentages may not total 100% due to missing data.
*Mean age represents mean age in 2013.
†Distribution of insurance types represents last known insurance for all patients.
ADG, Aggregated Diagnosis Groups (Johns Hopkins); DO, Doctor of Osteopathic Medicine; MD, Doctor of Medicine.

Table 2 Eligible providers' EHR behavior

Metric	All years*	Median per provider (25th percentile–75th percentile) 2010	2011	2012	Jan–Jun 2013	p for trend	Jan–Jun 2013 only: MU providers	Non-MU providers	p value
No. providers	112	55	70	90	87		41†	46	
Data access									
% Of encounters per provider with									
Problem list marked as reviewed	9.4 (3.8–15.5)	5.1 (0.9–19.9)	7.5 (1.6–13.7)	7.4 (1.9–11.3)	7.5 (4.2–14.6)	0.72	6.0 (3.7–10.0)	10.7 (6.8–20.7)	0.002
Medication list marked as reviewed	5.8 (2.4–15.5)	2.3 (0.1–10.9)	4.0 (0.7–11.1)	4.9 (1.6–10.4)	8.9 (4.5–15.3)	0.05	8.2 (3.6–11.0)	10.4 (5.2–17.1)	0.09
Allergies marked as reviewed	81.0 (73.3–85.1)	78.9 (71.6–82.8)	82.3 (75.2–86.9)	81.9 (73.0–87.6)	81.1 (73.7–86.0)	0.28	81.3 (73.3–86.6)	80.5 (73.7–85.9)	0.93
Problem list altered	23.6 (17.1–30.0)	19.5 (14.7–24.1)	23.0 (17.1–27.7)	22.2 (16.6–29.7)	25.0 (17.3–33.2)	0.002	22.8 (16.9–33.3)	27.7 (20.6–32.3)	0.39
≥1 problem added	23.6 (17.1–30.0)	19.5 (14.7–24.1)	23.0 (17.1–27.7)	22.2 (16.6–29.7)	25.0 (17.3–32.8)	0.002	22.7 (16.9–32.8)	27.6 (20.6–32.3)	0.39
≥1 problem replaced	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.07	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.75
≥1 problem dropped	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.39	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.36
Days to review laboratory results	4.4 (2.9–7.4)	4.9 (3.0–8.5)	5.3 (3.0–9.8)	4.1 (2.8–7.5)	4.0 (2.5–5.9)	0.01	4.0 (2.4–5.8)	4.0 (2.5–5.9)	0.59
Clinical decision support measures									
% Best-practice alerts 'accepted'	17.7 (9.4–24.6)	18.6 (13.0–27.0)	19.6 (12.6–30.0)	20.9 (10.7–30.0)	18.6 (10.4–26.0)	0.91	22.5 (13.8–28.4)	16.7 (8.6–24.4)	0.03
% Prescribing alerts 'accepted'	0.1 (0.0–0.2)	0.0 (0.0–0.3)	0.0 (0.0–0.3)	0.0 (0.0–0.2)	0.0 (0.0–0.1)	0.11	0.0 (0.0–0.1)	0.0 (0.0–0.1)	0.91
Drug–allergy alerts	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	NT	0.0 (0.0–0.0)	0.0 (0.0–0.0)	NT
Drug–disease alerts	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	NT	0.0 (0.0–0.0)	0.0 (0.0–0.0)	NT
Drug–drug interaction alerts	0.1 (0.0–0.3)	0.0 (0.0–0.4)	0.0 (0.0–0.4)	0.0 (0.0–0.2)	0.0 (0.0–0.2)	NT	0.0 (0.0–0.1)	0.0 (0.0–0.2)	NT
Drug–food alerts	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	NT	0.0 (0.0–0.0)	0.0 (0.0–0.0)	NT
Ordering									
Laboratory tests ordered per encounter	1.0 (0.7–1.4)	0.7 (0.4–0.9)	0.8 (0.5–1.1)	1.2 (0.8–1.5)	1.2 (0.9–1.6)	<0.001	1.2 (0.7–1.6)	1.3 (0.9–1.6)	0.44
Number of order set uses per encounter	1.1 (0.8–1.4)	1.1 (0.9–1.4)	1.1 (0.9–1.6)	1.1 (0.8–1.5)	1.1 (0.8–1.3)	0.61	1.1 (0.8–1.3)	1.0 (0.8–1.3)	0.36
Care coordination measures									
% Encounters with after-visit summary printed	59.2 (26.9–83.6)	10.8 (3.2–33.5)	47.8 (21.3–74.2)	73.1 (28.6–90.9)	89.5 (67.1–96.0)	<0.001	94.9 (89.5–97.6)	72.7 (37.0–90.8)	<0.001
Patients messaged/month	1 (0–3)	1 (0–3)	1 (1–4)	1 (0–4)	1 (0–4)	0.53	1 (0–5)	1 (0–3)	0.16
Provider-initiated messages/month	1 (0–4)	2 (0–4)	1 (1–5)	1 (0–4)	1 (0–4)	0.46	1 (0–6)	1 (0–3)	0.20
% Patient-initiated messages responded to	100 (98–100)	100 (100–100)	100 (99–100)	100 (98–100)	100 (99–100)	0.12	100 (94–100)	100 (100–100)	0.35
% Of patient panel using portal during year	6.9 (3.6–13.2)	11.3 (6.4–26.1)	10.8 (5.2–19.9)	8.2 (4.0–14.3)	6.0 (2.8–11.5)	<0.001‡	6.8 (3.8–14.5)	4.9 (2.5–11.3)	0.14
% Of portal users using portal ≥15 times/year	57.6 (51.6–64.7)	55.3 (50.9–65.6)	59.8 (53.2–67.1)	64.4 (54.3–72.1)	54.7 (43.8–66.0)	0.001‡	50.0 (41.8–58.8)	57.2 (48.2–66.7)	0.08
Population/panel level									
Viewed panel-level reports (median number of views)	0.0 (0.0–3.8)	NA	0.0 (0.0–5.0)	0.0 (0.0–5.0)	0.0 (0.0–1.0)	0.43	0.0 (0.0–3.0)	0.0 (0.0–0.0)	0.01
MU objective measures§									
Core									
Core 1—CPOE	100 (99–100)	NA	100 (99–100)	100 (99–100)	100 (99–100)	DNC	100 (100–100)	100 (99–100)	0.42
Core 3—Problem list	97 (95–98)	NA	96 (95–97)	97 (95–98)	98 (96–99)	0.02	98 (97–99)	98 (95–99)	0.51
Core 4—E-prescribing	68 (40–84)	NA	71 (44–84)	75 (46–87)	76 (50–87)	0.26	77 (64–88)	68 (37–86)	0.03
Core 5—Medication list	96 (93–97)	NA	95 (93–97)	95 (94–97)	97 (95–98)	0.01	97 (95–98)	97 (95–98)	0.67
Core 6—Allergy list	99 (99–100)	NA	99 (99–100)	99 (99–100)	99 (99–100)	0.17	99 (99–100)	99 (99–100)	0.44
Core 7—Demographics recorded	94 (91–97)	NA	94 (91–96)	94 (91–96)	95 (91–97)	0.43	96 (93–97)	95 (91–97)	0.94
Core 8—Vitals recorded	97 (96–98)	NA	97 (96–98)	98 (97–98)	98 (97–98)	<0.001	98 (97–98)	98 (97–98)	0.79
Core 9—Smoking status recorded	98 (97–99)	NA	98 (96–99)	99 (98–99)	99 (98–99)	0.75	99 (98–99)	98 (97–99)	0.36

Continued

Table 2 Continued

Metric	All years*	Median per provider (25th percentile–75th percentile)					Jan–Jun 2013 only		
No. providers	112	2010 55	2011 70	2012 90	Jan–Jun 2013 87	p for trend	MU providers 41†	Non-MU providers 46	p value
Core 11—Patient copy of health information	100 (100–100)	NA	–¶	–¶	100 (100–100)	NT	100 (100–100)	100 (100–100)	>0.99
Core 12—After-visit summary	63 (33–78)	NA	42 (28–64)	64 (31–81)	72 (49–84)	<0.001	77 (66–85)	65 (38–81)	0.01
Menu									
Menu 7—Timely access to data for patients	83 (76–90)	NA	73 (70–79)	81 (77–87)	91 (84–98)	<0.001	85 (83–91)	95 (88–100)	<0.001
Menu 8—Patient education	65 (57–75)	NA	56 (47–66)	69 (60–79)	73 (64–81)	<0.001	77 (69–85)	70 (60–75)	0.001
Menu 9—Medication reconciliation	15 (8–29)	NA	10 (3–19)	13 (5–27)	21 (11–38)	<0.001	18 (9–35)	25 (13–48)	0.18
Menu 10—Summary of care at transitions	77 (36–91)	NA	51 (25–78)	80 (39–94)	85 (58–97)	<0.001	92 (80–97)	70 (34–95)	0.01

For definitions of 'accepted' for best-practice alerts and prescribing alerts, please see the Methods section.
*Pooled results across all years are reported, except for panel-level reports and MU measures, for which weighted averages are reported.
†Although 42 providers attested at the end of 2012, one did not contribute data to the 2013 data set and is excluded from this comparison.
‡Comparisons do not include 2013 data as half-year results would not be expected to be similar to full-year results.
§Percentage of patients for whom metric was met, except for e-prescribing metric, which is calculated as a percentage of prescriptions. Core 2, 10, and 13 were accomplished at the organizational level and are not included as they do not vary by provider.
¶No provider had data for this measure in 2011 or 2012.
CPOE, computerized provider order entry; DNC, (model) did not converge; EHR, electronic health record; MU, meaningful use; NA, not available; NT, not tested.

after 42 IFH providers attested for MU in 2012, but the novel metrics developed for the current study showed even greater variability, presumably because there were no incentives associated with them. Some of the EHR functionalities measured here showed modest changes over the 3.5 years of this study. While some were directly linked to MU incentives (eg, the increase in use of after-visit summaries), in other cases there was no direct link (such as the drop in the number of days required to review laboratory test results).

One implication of these findings is that because individual-level use of EHR features to accomplish tasks varies, capturing this variation may add value to research on the quality and cost effects of health information technology.[10] One approach to examining individual-level use of EHRs is to employ the MU 'objective measures.' The stage I objective measures may, however, be too basic to use in a nuanced examination of provider behavior. For example, providers can achieve the MU problem-list metric by ensuring that patients have a single problem listed, but do not need to show that the problem list is complete or up-to-date. Similarly, providers can achieve the MU decision-support metric by demonstrating that the decision support functionality is enabled, but do not need to show that they have accepted any of the alerts.

Another approach to capturing and evaluating individual-level use of EHRs is by administering surveys[8 9 12] asking users to characterize the availability and use of EHR features. This approach allows a more granular analysis of the components of the EHR that might influence care. Poon *et al* employed a survey approach to demonstrate that EHR use as a binary factor is not associated with ambulatory physician quality performance, whereas use of specific EHR features was associated with significantly better quality scores.[8] Somewhat similarly, Amarasingham and colleagues used a survey tool to identify specific elements of the EHR associated with hospital quality and costs.[9] A very rich approach to assessing EHR use patterns was developed by Lanham and colleagues, who used interviews and direct observations to identify patterns of high, medium, and low EHR usage.[14]

Our approach is novel in that we employed EHR data to measure EHR use at the provider level, an approach that is potentially scalable. In addition, we attempted to develop sample metrics that might be expected to be associated with quality effects. For example, rather than recording only whether the patient problem list was populated, we captured whether the list was updated from encounter to encounter. Data capture directly from health IT systems has previously been explored for domain-specific studies, especially in CDS.[15–17 22 31 32] However, our goal was to develop sample metrics in a variety of different domains, so that alert override rates could be placed in the context of the larger issue of use of the EHR.

It is important to note that the metrics developed here should be interpreted as measures of intensity of use of the EHR, not necessarily as measures of better use of the EHR. For example, the electronic laboratory ordering metric (which increased significantly over time) should be seen as a measure of intensity of EHR use, not as a quality indicator. The high alert override rate also cannot necessarily be interpreted as an indicator of good or bad use of the EHR as it is consistent with the high override rates reported in an array of studies on contemporary CDS systems; [15–17 22 31 32] alert overrides can be attributed to factors ranging from alert fatigue to lack of clinical relevance in a particular case, and substantial proportions are found to be clinically appropriate on review.[33] The alert override rate is also dependent upon the severity threshold of the

Table 3 Effect of alert frequency on alert acceptance rate, by alert type

Clinical decision support measure	Median alerts received per provider per year (Q1–Q3)	Number of providers	Change in alert acceptance rate for every 100 additional alerts received: Bivariate model, Incidence rate ratio (95% CI)	Bivariate model, p Value	Multivariable (adjusted) model, Incidence rate ratio (95% CI)	Multivariable (adjusted) model, p Value
Best-practice alert type						
Total	*4227 (2057–6897)*	*112*	*0.997 (0.995 to 0.999)*	*<0.001*	*0.991 (0.987 to 0.994)*	*<0.001*
Pediatric immunization	501 (243–951)	112	0.90 (0.88 to 0.92)	<0.001	0.85 (0.82 to 0.88)	<0.001
Diabetes	428 (181–799)	111	1.00 (0.99 to 1.01)	0.74	1.01 (0.99 to 1.03)	0.38
Influenza	491 (219–752)	112	0.97 (0.95 to 0.99)	0.007	0.90 (0.87 to 0.95)	<0.001
Depression	344 (154–575)	112	0.98 (0.95 to 1.01)	0.18	0.98 (0.92 to 1.05)	0.57
Cervical cancer screening	237 (113–389)	109	0.79 (0.74 to 0.84)	<0.001	0.69 (0.63 to 0.76)	<0.001
Adult overweight/obesity	198 (92–408)	112	0.95 (0.86 to 1.05)	0.30	1.09 (0.97 to 1.23)	0.12
Tobacco	171 (71–317)	111	0.73 (0.67 to 0.80)	<0.001	0.59 (0.52 to 0.68)	<0.001
Hyperlipidemia	120 (53–252)	108	0.87 (0.82 to 0.92)	<0.001	0.94 (0.87 to 1.02)	0.16
Colorectal cancer screening	119 (61–236)	108	0.79 (0.73 to 0.82)	<0.001	0.78 (0.72 to 0.84)	<0.001
Pediatric overweight/obesity	71 (16–192)	105	0.78 (0.71 to 0.86)	<0.001	0.64 (0.56 to 0.74)	<0.001
Hypertension	57 (23–110)	108	0.38 (0.26 to 0.55)	<0.001	0.48 (0.29 to 0.80)	0.01
Breast cancer screening	51 (27–91)	108	0.43 (0.36 to 0.50)	<0.001	0.40 (0.33 to 0.49)	<0.001
Pneumonia	40 (16–83)	85	0.58 (0.51 to 0.67)	<0.001	0.58 (0.50 to 0.68)	<0.001
Asthma	24 (6–117)	112	0.68 (0.64 to 0.73)	<0.001	0.64 (0.60 to 0.70)	<0.001
Chlamydia	11 (4–24)	33	0.67 (0.10 to 4.48)	0.23	–	–
Alcoholism	11 (2–48)	23	0.55 (0.44 to 0.68)	<0.001	0.51 (0.37 to 0.70)	<0.001
Anticoagulation	7 (3–16)	92	0.01 (0.004 to 0.04)	<0.001	0.04 (0.01 to 0.13)	<0.001
All others	1091 (591–1879)	112	0.97 (0.97 to 0.98)	<0.001	0.94 (0.92 to 0.96)	<0.001
Drug alert type						
Total	*1527 (595–2568)*	*112*	*0.98 (0.96 to 0.99)*	*<0.001*	*0.99 (0.97 to 1.00)*	*0.12*
Drug–drug	931 (399–1724)	112	0.97 (0.95 to 0.99)	<0.001	0.98 (0.96 to 1.00)	0.11
Drug–food	482 (192–797)	111	0.80 (0.74 to 0.87)	<0.001	0.84 (0.70 to 1.01)	0.06
Drug–allergy	25 (10–53)	107	0.16 (0.06 to 0.39)	0.002	0.35 (0.12 to 1.01)	0.05
Drug–disease	4 (2–9)	97	–	–	–	–

Totals are italicized. Bivariate generalized linear mixed model (GLMM) was adjusted for clustering within provider and was weighted by year. Multivariable GLMM additionally included MU attestation in 2012, provider gender, degree (MD/DO vs APN), panel size, average ADG of panel. Cells marked with an ndash represent models that did not converge due to insufficient sample size. The incidence rate ratios (IRR) quantify the increase or decrease in the percentage of alerts accepted for every 100 additional alerts received; an IRR of 0.95 indicates that an alert type was accepted 5% less frequently for every additional 100 times it fired.
ADG, Aggregated Diagnosis Groups; APN, Advanced Practice Nurse; DO, Doctor of Osteopathic Medicine; MD, Doctor of Medicine; MU, meaningful use.

alerts shown to providers, which is generally determined at the institution level; unfortunately, we did not have severity or drug-class data on the alerts. Use of the patient portal is another metric that has not been definitively linked to improved quality.[34]

The provider-level variability observed here is most likely caused by a wide variety of factors at the product, organizational, provider, and patient levels. At the product level, system users reported that the panel-level view function in the EHR was rarely employed because it was difficult to use and training on this feature was not offered throughout the network. At the organizational level, allergies were most likely marked as reviewed more frequently than other elements of the record because the allergy review was typically performed by a medical assistant during patient intake, whereas other components (such as medications and problem lists) were reviewed by the provider during the encounter. Furthermore, during the study timeframe, new clinics joined the network and underwent EHR rollouts, resulting in subsets of providers who had lower familiarity with the EHR, as well as patients being newly introduced to the electronic patient portal. Although the EHR product was standardized across clinics, site-to-site staffing differences (not captured in our data) may have influenced EHR workflow. Providers who attested for MU, as expected, had higher scores on most MU measures that were not already affected by ceiling effects. However, they differed from the non-MU providers on only a few of the non-MU measures, suggesting the possibility that they were responding to specific incentives rather than demonstrating overall higher intensity of EHR use.

At the provider level, the data showed negative correlations between alert frequency and alert acceptance, both overall and by type. This finding is compatible with the explanation that providers are more likely to accept recommendations about conditions that they see less frequently and therefore are less familiar with. However, the findings are also compatible with an explanation of alert fatigue.[32 35–37]

Limitations

The metrics here were developed de novo for the current study and have not been validated with other data sets or demonstrated to be associated with healthcare quality. Limited information was available on the providers themselves in order to maintain confidentiality. The observed variability in usage may have been linked with other variables that we were unable to capture for the current analysis, including provider age and the provider's practice site at each time point in the study. A

number of patient-level characteristics are likely associated with these metrics, but patient-level analyses are outside the scope of the current study. This study was conducted at a single healthcare organization using a particular commercial product and may not be generalizable to other types of organizations or products.

Conclusions and implications

EHRs, like other information technologies, cannot be expected to cause the intended effects on healthcare delivery unless their features are regularly used.[10 38] Primary care providers at a single center using the same EHR product varied in their use of available functions of the EHR. This suggests that individual-level measures of usage may add value to future research on quality and cost outcomes of EHR use, even though more work remains to be done on explanations of this variability.

Acknowledgements This study was conducted as part of the work of the Health Information Technology Evaluation Collaborative (HITEC), an academic consortium recognized by the state of New York as evaluators for state-funded information technology initiatives. The authors would also like to acknowledge the valuable contributions of three anonymous reviewers during the review process.

Collaborators Health Information Technology Collaborative (HITEC).

Contributors JSA conceived and designed the study, contributed to data analysis, interpreted results, and drafted and revised the paper. LMK contributed to study design and data interpretation and revised the paper. AE conducted statistical analysis and data management. SN assisted with data acquisition and data interpretation. DMS contributed to study design and data interpretation. DH assisted with data acquisition and data interpretation. RK contributed to study conception and design, data interpretation, and revisions of the paper.

Funding This study was supported by the New York State Department of Health (contract C025877).

Competing interests None.

Ethics approval Weill Cornell IRB and the Institute for Family Health IRB approved this study.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Qualified researchers may apply to the Health Information Technology Evaluation Collaborative for access to data.

REFERENCES

1 Blumenthal D, Tavenner M. The "meaningful use" regulation for electronic health records. *N Engl J Med* 2010;363:501–4.
2 Kaushal R, Shojania KG, Bates DW. Effects of computerized physician order entry and clinical decision support systems on medication safety: a systematic review. *Arch Intern Med* 2003;163:1409–16.
3 Chaudhry B, Jerome W, Shinyi W, *et al*. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. *Ann Intern Med* 2006;144:742–52.
4 Han YY, Carcillo JA, Venkataraman ST, *et al*. Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system. *Pediatrics* 2005;116:1506–12.
5 Koppel R, Metlay JP, Cohen A, *et al*. Role of computerized physician order entry systems in facilitating medication errors. *JAMA* 2005;293:1197–203.
6 Upperman JS, Staley P, Friend K, *et al*. The impact of hospitalwide computerized physician order entry on medical errors in a pediatric hospital. *J Pediatr Surg* 2005;40:57–9.
7 Del Beccaro MA, Jeffries HE, Eisenberg MA, *et al*. Computerized provider order entry implementation: no association with increased mortality rates in an intensive care unit. *Pediatrics* 2006;118:290–5.
8 Poon EG, Wright A, Simon SR, *et al*. Relationship between use of electronic health record features and health care quality: results of a statewide survey. *Med Care* 2010;48:203–9.
9 Amarasingham R, Plantinga L, Diener-West M, *et al*. Clinical information technologies and inpatient outcomes: a multiple hospital study. *Arch Intern Med* 2009;169:108–14.
10 Ancker JS, Kern LM, Abramson E, *et al*. The triangle model for evaluating the effect of health information technology on healthcare quality and safety. *JAMIA* 2011;18:749–53.
11 Berg M, Aarts J, van der Lei J. ICT in health care: sociotechnical approaches. *Methods Inf Med* 2003;42:297–301.
12 Amarasingham R, Diener-West M, Weiner M, *et al*. Clinical information technology capabilities in four U.S. hospitals: testing a new structural performance measure. *Med Care* 2006;44:216–24.
13 Makam AN, Lanham HJ, Batchelor K, *et al*. Use and satisfaction with key functions of a common commercial electronic health record: a survey of primary care providers. *BMC Med Inform Decis Mak* 2013;13:86.
14 Lanham HJ, Sittig DF, Leykum LK, *et al*. Understanding differences in electronic health record (EHR) use: linking individual physicians' perceptions of uncertainty and EHR use patterns in ambulatory care. *J Am Med Inform Assoc* 2014;21:73–81.
15 Tamblyn R, Reidel K, Patel V. Physicians' response to computerised alerts for psychotropic drugs in older persons: a multilevel analysis of the associated alert, patient and physician characteristics. *BMJ Open* 2012;2:e001384.
16 van der Sijs H, Aarts J, Vulto A, *et al*. Overriding of drug safety alerts in computerized physician order entry. *J Am Med Inform Assoc* 2006;13:138–47.
17 Eslami S, Abu-Hanna A, de Keizer NF. Evaluation of outpatient computerized physician medication order entry systems: a systematic review. *J Am Med Inform Assoc*. 2007;14:400–6.
18 Ash J. Organizational factors that influence information technology diffusion in academic health sciences centers. *J Am Med Inform Assoc* 1997;4:102–11.
19 Holdren JP, Lander E, Varmus H. The President's Council of Advisors on Science and Technology (PCAST), eds. *Report to the President: Realizing the full potential of health information technology to improve health care for Americans: The path forward*. 2010.
20 Covell D, Uman G, Manning P. Information needs in office practice: are they being met? *Ann Intern Med* 1985;103:596–9.
21 Speedie SM, Park Y-T, Du J, *et al*. The impact of electronic health records on people with diabetes in three different emergency departments. *J Am Med Inform Assoc* 2014;21(e1):e71–7.
22 Wright A, Pang J, Feblowitz JC, *et al*. Improving completeness of electronic problem lists through clinical decision support: a randomized, controlled trial. *J Am Med Inform Assoc.* 2012.
23 Hartung DM, Hunt J, Siemienczuk J, *et al*. Clinical implications of an accurate problem list on heart failure treatment. *J Gen Intern Med* 2005;20:143–7.
24 Bates DW, Kuperman GJ, Wang S, *et al*. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. *J Am Med Inform Assoc* 2003;10:523–30.
25 Miller RA, Waitman LR, Chen S, *et al*. The anatomy of decision support during inpatient care provider order entry (CPOE): empirical observations from a decade of CPOE experience at Vanderbilt. *J Biomed Inform* 2005;38:469–85.
26 Kuperman GJ, Bobb A, Payne TH, *et al*. Medication-related clinical decision support in computerized provider order entry systems: a review. *J Am Med Inform Assoc* 2007;14:29–40.
27 Kaushal R, Kern LM, Barron Y, *et al*. Electronic prescribing improves medication safety in community-based office practices. *J Gen Intern Med* 2010;25:530–6.
28 Stead WW, Lin HS., eds. *Computational technology for effective health care: immediate steps and strategic directions*. Washington, DC: National Academies Press, 2009.
29 Anonymous. Stage 1 Eligible Professionals Attestation Worksheet. Washington, DC. 2014. http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Downloads/EP_Attestation_Stage1Worksheet_2014Edition.pdf
30 Weiner J, Starfield B, Steinwachs D, *et al*. Development and application of a population-oriented measure of ambulatory care case-mix. *Med Care* 1991;29:452–72.
31 Kuperman GJ, Teich JM, Tanasijevic MJ, *et al*. Improving response to critical laboratory results with automation: results of a randomized controlled trial. *J Am Med Inform Assoc* 1999;6:512–22.
32 Shah NR, Seger AC, Seger DL, *et al*. Improving acceptance of computerized prescribing alerts in ambulatory care. *J Am Med Inform Assoc* 2006;13:5–11.
33 Nanji KC, Slight SP, Seger DL, *et al*. Overrides of medication-related clinical decision support alerts in outpatients. *J Am Med Inform Assoc* 2014;21:487–91.
34 Ammenwerth E, Schnell-Inderst P, Hoerbst A. The impact of electronic patient portals on patient care: a systematic review of controlled trials. *J Med Internet Res* 2013;14:e162.
35 Ash JS, Sittig DF, Campbell EM, *et al*. Some unintended consequences of clinical decision support systems. *AMIA Annu Symp Proc* 2007:26–30.
36 Kesselheim AS, Cresswell K, Phansalkar S, *et al*. Clinical decision support systems could be modified to reduce 'alert fatigue' while still minimizing the risk of litigation. *Health Aff (Millwood)* 2011;30:2310–7.
37 Xie M, Johnson K. Applying human factors research to alert-fatigue in e-prescribing. *AMIA Annu Symp Proc* 2007:1161.
38 DeLone WH, McLean ER. The Delone and McLean model of information systems success: a 10-year update. *J Manag Info Syst* 2003;19:9–30.