

The Development and Piloting of the Ambulatory Electronic Health Record Evaluation Tool: Lessons Learned

Zoe Co¹ A. Jay Holmgren² David C. Classen³ Lisa P. Newmark⁴ Diane L. Seger⁴ Jessica M. Cole⁵
Barbara Pon⁶ Karen P. Zimmer⁷ David W. Bates^{1,4,8}

¹ Division of General Internal Medicine, Brigham and Women's Hospital, Boston, Massachusetts, United States

² Harvard Business School, Boston, Massachusetts, United States

³ Department of Clinical Epidemiology, University of Utah, Salt Lake City, Utah, United States

⁴ Clinical and Quality Analysis, Mass General Brigham, Somerville, Massachusetts, United States

⁵ Department of Internal Medicine, Division of Epidemiology, University of Utah, Salt Lake City, Utah, United States

⁶ Collaborative Healthcare Patient Safety Organization, Sacramento, California, United States

⁷ Department of Pediatrics, Thomas Jefferson University, Philadelphia, Pennsylvania, United States

⁸ Harvard Medical School, Boston, Massachusetts, United States

Address for correspondence Zoe Co, BS, Department of General Internal Medicine, Brigham and Women's Hospital, 1620 Tremont Street, 3rd Floor, Boston, MA 02120, United States (e-mail: zco@bwh.harvard.edu).

Appl Clin Inform 2021;12:153–163.

Abstract

Background Substantial research has been performed about the impact of computerized physician order entry on medication safety in the inpatient setting; however, relatively little has been done in ambulatory care, where most medications are prescribed.

Objective To outline the development and piloting process of the Ambulatory Electronic Health Record (EHR) Evaluation Tool and to report the quantitative and qualitative results from the pilot.

Methods The Ambulatory EHR Evaluation Tool closely mirrors the inpatient version of the tool, which is administered by The Leapfrog Group. The tool was piloted with seven clinics in the United States, each using a different EHR. The tool consists of a medication safety test and a medication reconciliation module. For the medication test, clinics entered test patients and associated test orders into their EHR and recorded any decision support they received. An overall percentage score of unsafe orders detected, and order category scores were provided to clinics. For the medication reconciliation module, clinics demonstrated how their EHR electronically detected discrepancies between two medication lists.

Results For the medication safety test, the clinics correctly alerted on 54.6% of unsafe medication orders. Clinics scored highest in the drug allergy (100%) and drug–drug interaction (89.3%) categories. Lower scoring categories included drug age (39.3%) and therapeutic duplication (39.3%). None of the clinics alerted for the drug laboratory or drug monitoring orders. In the medication reconciliation module, three (42.8%) clinics had an EHR-based medication reconciliation function; however, only one of those clinics could demonstrate it during the pilot.

Keywords

- ▶ electronic health record
- ▶ computerized physician order entry
- ▶ ambulatory care
- ▶ patient safety
- ▶ quality of care
- ▶ medication safety
- ▶ medication reconciliation

received
August 11, 2020
accepted after revision
December 16, 2020

© 2021. Thieme. All rights reserved.
Georg Thieme Verlag KG,
Rüdigerstraße 14,
70469 Stuttgart, Germany

DOI <https://doi.org/10.1055/s-0041-1722917>.
ISSN 1869-0327.

Conclusion Clinics struggled in areas of advanced decision support such as drug age, drug laboratory, and drug monitoring. Most clinics did not have an EHR-based medication reconciliation function and this process was dependent on accessing patients' medication lists. Wider use of this tool could improve outpatient medication safety and can inform vendors about areas of improvement.

Background and Significance

The Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009 provided approximately 40 billion dollars in federal funds for hospitals and physician offices to adopt electronic health records (EHRs).¹ One of the motivating factors behind this public investment, was that EHRs were theorized to improve quality and reduce medical errors. One mechanism for improving quality is computerized physician order entry (CPOE) and the ability of the EHR to deliver clinical decision support (CDS) at the point of care, which can prevent mistakes before harming a patient.^{2,3}

CDS can be delivered interruptively or non-interruptively and can appear during or after medication ordering. Facilities can customize their CDS to a substantial degree in terms of which alerts to turn on and, sometimes, when these alerts appear. This customization can impact the quality and safety of care, so that if too few alerts are delivered, adverse drug events (ADEs) may occur, and if there is overalerting, providers may develop alert fatigue, potentially causing them to miss important alerts.⁴ While many studies have found that CPOE with decision support has improved medication safety,^{5,6} the degree of organization-level configurability has resulted in significant variation in safety performance, even between organizations using the same EHR system.⁷⁻⁹ It is therefore critical to study CPOE safety performance at the organizational level, rather than simply relying on vendor-wide assessments.

Substantial research has been done in the inpatient setting on the impact of CDS on medication safety; however, relatively little has been done in ambulatory care, where most medications are prescribed. One early study by Kaushal et al¹⁰ found that using an electronic prescribing system decreased medication errors sevenfold across several community-based practices, while those that did not have this system, continued to have a high error rate. In another study, Gandhi et al¹¹ concluded that advanced drug ordering decision support capabilities such as dose and frequency checking can prevent ADEs. More recently, the negative impact of EHRs on clinician well-being has become more apparent, and decision support systems that bombard clinicians with hundreds of pop-up alerts per day may be a significant contributor to information technology (IT)-induced burnout.¹² Even more distressingly, studies have found clinicians who receive too many alerts may suffer from "alert fatigue" and begin to disregard even the most important CDS alerts.¹³ Provider organizations must balance their CDS to alert clinicians to potential ADEs without bombarding them with too many alerts.¹⁴

To address these critical issues regarding EHR safety performance in outpatient settings, the Ambulatory EHR Evaluation Tool was developed and piloted with seven outpatient clinics to assess outpatient clinic EHR's ability to assess common and serious medication-related errors. The tool currently consists of two sections, a medication safety test and a medication reconciliation module, and will be further developed to evaluate other safety domains such as usability. The medication safety test closely mirrors the inpatient version of the tool, which is administered by The Leapfrog Group and has been extensively validated.^{7,8,14-17} Clinics receive test patients and associated medication test orders to enter into their EHR using CPOE, and record any decision support they receive. Next, clinics receive an overall percentage score of unsafe orders detected and individual order category scores. The medication reconciliation module assesses the ability of clinics' EHRs to electronically detect medication discrepancies by having clinics demonstrate how they perform this process at their facility. With this tool, outpatient clinics can identify some of the gaps in their EHR implementation, and researchers and policy-makers can use the aggregated results to gain a better understanding of the state of EHR medication safety.

Objectives

In this article, the development and piloting process of the Ambulatory EHR Evaluation Tool is presented, along with the quantitative and qualitative results from the seven-clinic pilot.

Methods

Development of the Ambulatory EHR Evaluation Tool

The Ambulatory EHR Evaluation Tool was developed in collaboration with The University of Utah, Brigham and Women's Hospital, the Collaborative Healthcare Patient Safety Organization, and the Institute of Health Improvement. Researchers from these institutions collaborated in developing the tool and advised on the piloting process.

Medication Safety Test

The content of the medication safety test was derived from the inpatient version of this tool, which is administered by The Leapfrog Group and is endorsed by the National Quality Forum as part of their "Safe Practices for Better Healthcare Report."¹⁸ A group of experts specializing in ADEs and CDS within CPOE systems created the content of the inpatient tool.^{8,16} The test patients and medication test orders were created based on real-world cases, where patients were either severely injured or died from preventable ADEs.¹⁶ The medication test orders

within the inpatient test covers basic and advanced decision support features.¹⁹

The content was also based off of a study by Gandhi et al,²⁰ that identified common types of ADEs that occurred in the ambulatory setting. Some of the preventable ADEs that they determined were the result of an incorrect dose or the medication was taken at an incorrect frequency. In addition, they found that decision support features such as allergy-checking and drug–drug interaction checking could have prevented 35% of the preventable ADEs identified in their study. Given these findings, these types of alerts were also added to the content library.

With these resources, a research pharmacist first reviewed the inpatient content and removed orders that were not applicable to the outpatient setting. In this case, the drug route category from the inpatient tool was removed because most orders in this category were injectables. To replace this category, a new category called “drug pregnancy” was added, which included medication orders that are contraindicated in pregnant patients. The other categories in the inpatient test were still applicable to outpatient settings and were kept as part of the content. With these changes, the medication safety test consisted of 10 order checking categories that assess basic and advanced decision support features (→Table 1). Order categories that were categorized

as basic decision support² included drug allergy, single dosing, therapeutic duplication, and drug–drug interaction. Areas of advanced decision support² included daily cumulative dosing (drug dose [daily]), drug age, drug laboratory, drug monitoring, drug diagnosis, and drug pregnancy.

The content also included two subcategories: fatal and nuisance orders. Fatal orders have the potential to cause serious injury or death to patients. These types of orders were included in the drug dose (daily), drug–drug interaction, and drug allergy categories. An example of a fatal drug dose (daily) order is prescribing prasugrel 60 mg orally daily. The nuisance orders tested whether a clinic’s decision support system overalerts by including several orders that are low priority and should not generate decision support warnings. These orders, derived from a study by Phansalkar et al,²¹ can be in multiple order categories (drug–drug and therapeutic duplication) but none would cause patient harm. An example of a nuisance order is prescribing furosemide 20 mg orally daily with digoxin 0.25 mg orally daily, which often triggers alerts when it should not, which can burden physicians.²¹

Once the research pharmacist finished reviewing the content, it was sent to an expert panel that consisted of physicians and pharmacists, who reviewed the content and made additional suggestions; specifically, in the drug monitoring, drug diagnosis, and drug allergy categories. These suggestions were

Table 1 The order checking categories in the medication safety test, covering both basic and advanced decision support features,² with examples from the actual test

Basic decision support		
Order category	Description	Example
Drug allergy	Medication is one for which a patient allergy has been documented	Penicillin prescribed for patient with documented penicillin allergy
Drug dose (single)	Specified dose of medication exceeds the safe range for a single dose	Tenfold overdose of digoxin
Therapeutic duplication	Medication combinations overlap therapeutically (same agent or class)	Use of clonazepam and lorazepam together
Drug–drug interaction	Medication order pairs that result in a known harmful interaction when used in combination	Concurrent use of sumatriptan and phenelzine
Advanced decision support		
Order category	Description	Example
Drug dose (daily)	Cumulative dose for medication exceeds the safe range for daily dose	Ordering ibuprofen regular dose every three hours
Drug age	Medication dose inappropriate/contraindicated based on patient’s age	Prescribing diazepam for a patient over 65 years old
Drug laboratory	Medication dose inappropriate/contraindicated based on documented laboratory results (including renal status)	Use of nitrofurantoin in patient with severe renal failure
Drug monitoring	Medication for which the standard of care includes subsequent monitoring of drug level or laboratory value to avoid harm	Prompt to monitor drug levels when ordering digoxin or monitor INR/PT when ordering warfarin
Drug diagnosis	Medication dose inappropriate/contraindicated based on document diagnosis	Prescribing a nonspecific beta-blocker for patient with asthma
Drug pregnancy	Medication inappropriate/contraindicated in pregnant patients	Prescribing atorvastatin for a pregnant patient

Abbreviations: INR, international normalized ratio; PT, prothrombin time.

then sent back to the research pharmacist, and additional content changes were made. In total, there were two rounds of review by the expert panel before a finalized version of the content was created for the pilot test.

Medication Reconciliation Module

For the medication reconciliation module, experts from Brigham and Women's Hospital created test scenarios based on actual cases that reflected situations where medication reconciliation was required. All the test cases involved a patient that was recently discharged from the hospital and was returning to an outpatient clinic for a follow-up visit. Two medication lists were created for each patient: the most recent ambulatory medication list prior to the hospital admission and the discharge medication list from the hospital. These medication lists contained at least one of the following discrepancies: medication(s) that are longer on the medication list, addition of a new medication, or a change in the dose of an existing medication. Once these cases were created, the research team edited them for clarity and translated them into a testable format.

Testing Methodology of the Ambulatory EHR Evaluation Tool

The testing methodology of the medication safety test mirrors that of the inpatient version of the tool, in that it simulates the action of a physician entering orders for a patient to assess the EHR's performance against common and serious prescriber errors.^{8,15,16} For the outpatient version of the tool, the test simulates physicians prescribing medications to their patients. Each clinic received a set of test patients and associated test orders that a physician entered into their EHR using CPOE. For each patient, demographic and clinical information such as allergies, diagnoses, and relevant laboratory values were provided (→ **Supplementary Table S1**, available in the online version). While physicians entered each test order, they recorded any advice or information they received on the Orders and Observation Sheet (→ **Supplementary Table S2**, available in the online version).

Once finished with the medication safety test, clinics received an overall percentage score of unsafe orders detected, as well as percentage scores for each order category. These percentage scores were calculated by dividing the number of orders correctly alerted on by the total number of electronically orderable orders. For example, if a clinic did not have a medication on their formulary, that test order was removed from their denominator and numerator. This indicates that the denominator for each clinic can vary. Since there is great variability in how alerts are delivered, there was no difference in scoring based on what form the CDS was delivered. For example, hard stops, non-interruptive alerts, guidance, messages, and other information were all scored the same.

For the fatal orders and nuisance orders, clinics were provided with the specific orders they missed during the test. This is so that adjustments can be made to their EHR system to avoid serious patient harm and to help prevent overalerting. Nuisance orders were reverse scored, in that a higher percentage score indicated that the clinic correctly

avoided alerting on many nuisance orders, while a lower percentage score indicated that the clinic alerted on several nuisance orders. As with the inpatient test, this first round of piloting did not factor in nuisance order scores into the calculation of the overall score.¹⁴

Also included in the test were normal and safe orders that were used to ensure that clinics were taking the test as intended. More specifically, they were used to discourage clinics from recording that they received alerts for every test order to potentially achieve a higher overall score. If a clinic alerted on more than two of these orders, their test was invalid.

For the medication reconciliation module, only one test scenario was used. There were three discrepancies between the two medication lists: medication(s) that are longer on the medication list, addition of a new medication, or a change in the dose of an existing medication. Clinics demonstrated how their facility would electronically reconcile these medication lists.

Pilot Process and Testing

To recruit clinics for the pilot, the 2017 Office of the National Coordinator database²² was used to identify the seven leading outpatient EHR vendors in the United States. Only seven vendors were chosen to participate in the pilot due to resource constraints. Once the vendors were identified, the research team asked those vendors to recommend clinics that would be interested in participating in the pilot. The research team then sent out a recruitment email that described the purpose of the tool and the general testing process. Once these clinics agreed to participate, the pilot was administered in three phases using a webinar and all sessions were recorded.

In Phase 1, the tool was introduced to the clinic, and a sample test was administered through a Portable Document Format (→ **Supplementary Tables S1** and **S2**, available in the online version). The sample test consisted of one test patient and two test orders (→ **Fig. 1**).

In Phase 2, both sections of the test were administered (→ **Fig. 1**). For the medication safety test, the same test was used for all seven clinics. Following the methodology used in the inpatient tool and to ensure the feasibility of this test, the pilot test only included 11 test patients and 48 medication test orders. In the inpatient tool, the average time to complete the test is 2 to 3 hours.⁸ Also, if the licensed prescriber at the clinic could customize the level of alerts they see, they were asked to set it to the "normal" level that most prescribers used across their practice. While the physician was entering the test orders, the research team documented the alerts they received, and paused the session when necessary to clarify with the clinic whether an alert was triggered. After this section was finished, the research staff compared their scores and if there was a disagreement, the staff rewatched the recording and consulted the principal investigators for guidance. In the medication reconciliation module, clinics demonstrated on the webinar, how discrepancies between medication lists are detected. For clinics that did not have an EHR-based medication reconciliation functionality, they were asked to describe how they identify discrepancies between medication lists. This module was not scored.

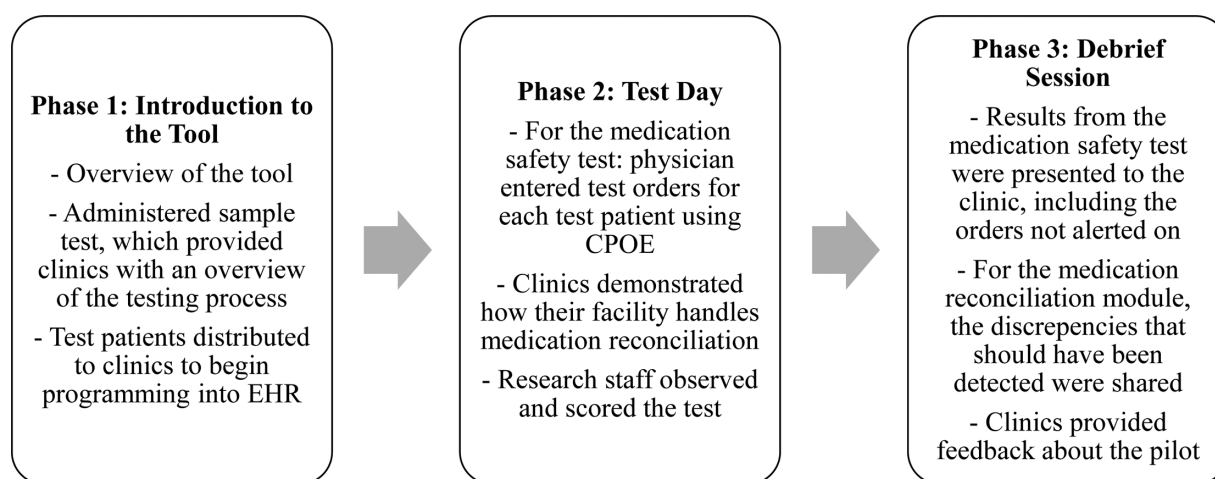


Fig. 1 The process used to pilot the Ambulatory Electronic Health Record (EHR) Evaluation Tool.

During Phase 3, a debrief session was held with each clinic individually to discuss their results from both sections of the tool. Clinics had the opportunity to provide suggestions for the test and provide comments about their testing experience (→Fig. 1).

Analysis

For the seven pilot clinics, clinic organizational characteristics such as service lines, volume, number of clinicians, payer distribution, value-based payment model participation, source of IT support, and what quality metrics they report, were acquired from a pretest survey administered to each clinic. Descriptive statistics including overall, nuisance order, and fatal order scores for each clinic, as well as percentage scores for each category were reported. Lastly, the qualitative results from the medication reconciliation module were presented.

The University of Utah's Institutional Review Board (00107070) deemed this study as nonhuman subject research. The Mass General Brigham Institutional Review Board also reviewed the study (Protocol #2018P001197) and determined that the Brigham and Women's Hospital component of the study was not human subject research.

Results

Clinical Organizational Characteristics

From September 2019 to December 2019, the Ambulatory EHR Evaluation Tool was piloted with seven outpatient clinics that represent the seven leading outpatient EHR vendors. The clinics were located in two regions in the United States, the Northeast and the West. Clinics of varying sizes and specialties were recruited (→Table 2). The majority (71%) of clinics were part of a health care system. Most (57%) clinics only provided primary care and the remainder of the clinics were multispecialty clinics. Patients at these clinics were covered by various types of insurance, with some patients mainly covered by Medicare or Medicaid and others covered by private insurance. Most clinics (71%) used multiple value-based payment models, while two clinics

(29%) used only one type. In terms of how and where clinics receive IT support, most clinics (71%) have in-house IT support staff; only two clinics (29%) received IT support from their EHR vendor. Next, clinics reported their quality metrics to several different organizations and all clinics at a minimum report to the Centers for Medicare and Medicaid Services. Lastly, all the clinics were required to report quality metrics, and which metrics they reported on, were decided by several groups. Some of these include their physician organization and the organizations to whom they report these metrics.

Medication Safety Test Performance

The mean overall percentage score for the medication safety test was 54.6%. Overall scores ranged from 37.5 to 80% (→Table 3). Clinics scored highest in the following categories: drug allergy (100%), drug–drug interaction (89.3%), drug pregnancy (75%), daily dosing (78.6%), and drug diagnosis (67.9%). Lower scoring categories included single dosing (57.1%), drug age (39.3%), and therapeutic duplication (39.3%). None of the clinics alerted on the drug laboratory and drug monitoring orders in the test (→Fig. 2). The average time it took to complete the medication safety test was 2 hours.

The mean fatal order score was 67.9% (→Table 3). All clinics alerted on at least half of the fatal orders in their test, and only one clinic alerted on all of them. The mean nuisance order score was 64.3%, and scores ranged from 0 to 100%.

Medication Reconciliation Module: Qualitative Results

For the medication reconciliation module, three (43%) clinics used an EHR-based medication reconciliation function that notified providers if there were any discrepancies between the medication lists; the provider would confirm that the medication list was updated. However, only one of the clinics could actually demonstrate this functionality in their EHR during Phase 2 of piloting. One clinic noted that their EHR system does not provide any CDS during this process. For the other clinics, patient's medication lists were stored in the EHR, but the actual comparison of medication lists was performed manually by a nurse or medical assistant.

Table 2 Demographic and office practice information collected from clinics in the pretest survey

Clinic	A	B	C	D	E	F	G	Percentages
Health care system or standalone clinic								
Part of health care system	Yes	No	No	Yes	Yes	Yes	Yes	71.4%
Standalone	No	Yes	Yes	No	No	No	No	28.6%
Types of services								
Primary Care	No	No	Yes	Yes	No	Yes	Yes	57.1%
Multispecialty	Yes	No	No	No	Yes	No	No	28.6%
Both	No	Yes	No	No	No	No	No	14.3%
Total number of visits each year								
Total	50,000	100,000	12,000	2,500	≥ 320,000	18,000	16,500	N/A
Number of physicians								
Primary care	100	60	2	20	70	5	2	N/A
Specialty	500	80	N/A	N/A	Unknown	N/A	N/A	N/A
Number of mid-level providers								
Total	≥ 100	40	1	2	90	1	2	N/A
Percentage of patient population covered by:								
Medicaid	25%	20%	8%	10%	15%	37%	30%	N/A
Medicare	50%	20%	25%	30%	15%	23%	20%	N/A
Dual eligible Medicare/Medicaid	N/A	10%	8%	2%	5%	0%	10%	N/A
Private insurance	25%	45%	55%	50%	60%	36%	50%	N/A
Uncompensated care	Unknown	5%	4%	8%	Less than 5%	40%	5%	N/A
Participation in value-based payment models								
Medicare ACO	Yes	Yes	No	No	Yes	Yes	Yes	71.4%
Private payer ACO	Yes	Yes	No	Yes	Yes	Yes	No	71.4%
Patient-centered medical home	Yes	No	Yes	Yes	Yes	Yes	No	57.1%
Other	N/A	N/A	N/A	N/A	Medicaid ACO	N/A	N/A	N/A
How clinic receives IT support								
From physician organization	Yes	No	Yes	Yes	No	No	No	42.9%
In-house IT support staff	No	Yes	Yes	No	Yes	Yes	Yes	71.4%
EHR vendor	No	Yes	Yes	No	Yes	No	No	42.9%

Table 2 (Continued)

Clinic	A	B	C	D	E	F	G	Percentages
Organization clinic reports quality metrics to								
Organizations	Medicare	CMS	CMS	CMS and commercial payers	CMS, BayCare, NCQA, BCBS, HNE, and many others	ACO, Medicare via EMR reporting service	CMS commercial payers, NCQA	N/A
Are quality metrics required to be reported?								
Yes	Based on input from quality leadership of our physician organization	Agreed to in our ACO contract	Based on importance to providers and risk reduction for patients	Physician organization decides for them	The organization they report to decides	These are required metrics	Report standard measures that are required for primary care	N/A

Abbreviations: ACO, Affordable Care Organization; BCBS, Blue Cross Blue Shield Association; CMS, Centers for Medicaid and Medicare Services; HNE, Health New England; IT, information technology; NCQA, National Committee for Quality Assurance.

Note: For clinics A and E, they report on their health care system, since they could not access the information for their individual clinic.

Discussion

The Ambulatory EHR Evaluation Tool was developed and piloted with seven outpatient clinics, each using one of the seven leading outpatient EHR vendor systems. There was significant performance variation; most clinics had basic decision features implemented, while advanced decision support—which likely delivers a large part of the safety benefit—was largely absent. The mean overall percentage score was 54.6%, indicating a little over half of the medication test orders were correctly alerted on. The mean fatal order and mean nuisance order scores were 67.9% and 64.3%, respectively. For the medication reconciliation module, most clinics did not electronically reconcile medication lists, and usually had a clinician identify discrepancies between medication lists.

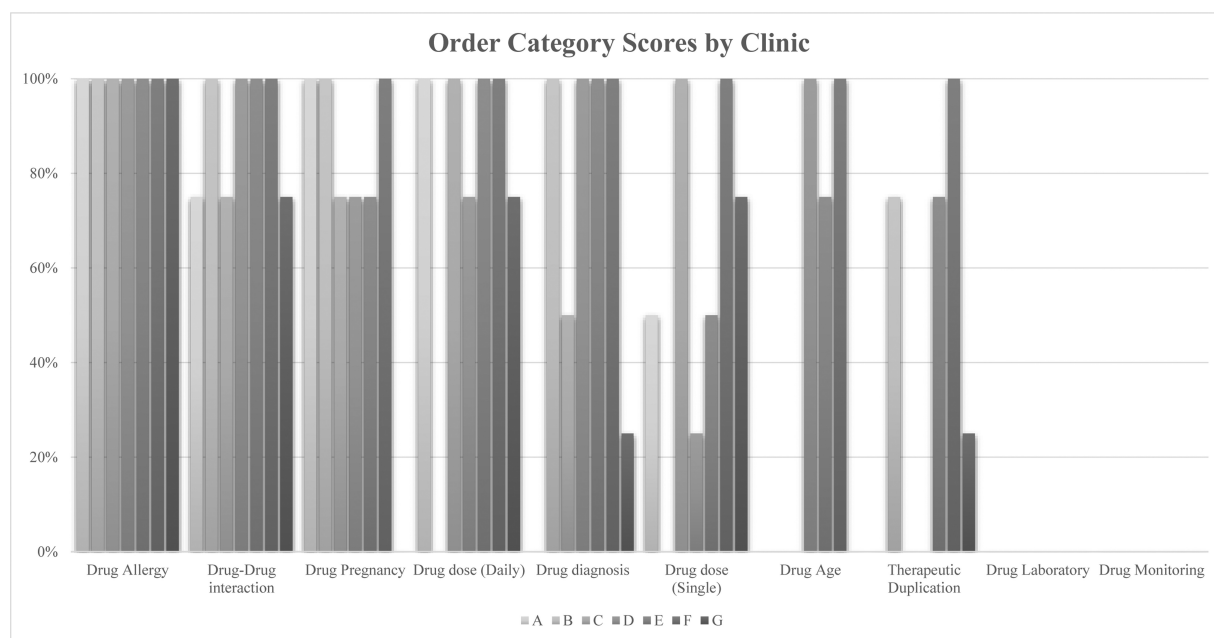
The results from the medication safety test provide a high-level overview of the types of alerts outpatient clinics have in place. In this evaluation, all the clinics had a mix of basic and advanced decision support features implemented in their systems, with clinics generally performing quite well for basic decision support categories such as drug allergy and drug–drug interaction. However, for other categories, a common theme emerged, in that certain types of alerts were completely absent. This was illustrated by Clinic B, where they scored 0% in the drug dose (daily) and drug dose (single) order categories, both of which carry substantial patient risk. This clinic confirmed that these alerts were turned off in their system. In addition, the drug age and therapeutic duplication categories had the most variability in performance across all the clinics, where several of them scored 0% in these categories, while other clinics correctly alerted for almost all those orders. One clinic commented that the drug age alerts were not turned in their system because their medication reference database does not require it. This is especially alarming since this category focuses on geriatric alerts. This also suggests that there may be a limit as to how much control outpatient clinics have over the customization of their EHR. Since most of the clinics had an in-house IT support staff, these results can be used to aid in customizing when and how alerts are triggered in these clinics' systems.^{23,24}

Another commonality between the clinics was that none of them alerted for the drug laboratory and drug monitoring orders. A scenario where a drug laboratory alert should appear is when nitrofurantoin is prescribed to a patient with severe renal failure. An example of a drug monitoring alert is a prompt to monitor lithium levels after starting a patient on lithium carbonate. All the clinics reported that these alerts are not turned on in their system, even though their EHR has this functionality, and these are some of the most dangerous clinical situations. One clinic reported that there is difficulty in obtaining the most updated lab results if the patient did not have their laboratory work done at a facility associated with their clinic. This suggests that there are barriers to obtaining the most updated laboratory information, and this can be a result of the lack of health data exchange between facilities.²⁵ There are also cases of clinics within the same health care system having different procedures for following up on abnormal laboratory results.²⁶ For clinics that did have updated lab

Table 3 The results from the seven-clinic pilot

Category	Clinic A %	Clinic B %	Clinic C %	Clinic D %	Clinic E %	Clinic F %	Clinic G %	Mean %
Drug allergy	100	100	100	100	100	100	100	100
Drug–drug interaction	75	100	75	100	100	100	75	89.3
Drug pregnancy	100	100	75	75	75	100	0	75
Drug dose (daily)	100	0	100	75	100	100	75	78.6
Drug diagnosis	0	100	50	100	100	100	25	67.9
Drug dose (single)	50	0	100	25	50	100	75	57.1
Drug age	0	0	0	100	75	100	0	39.3
Therapeutic duplication	0	75	0	0	75	100	25	39.3
Drug laboratory	0	0	0	0	0	0	0	0
Drug monitoring	0	0	0	0	0	0	0	0
Overall score	42.5	47.5	50	57.5	67.5	80	37.5	54.6
Fatal order score	50	50	75	75	75	100	50	67.9
Nuisance order score	100	50	100	100	25	0	75	64.3

Note: The mean overall percentage of unsafe orders detected was 54.6%. The mean fatal order score was 67.9% and the mean nuisance order score was 64.3%.

**Fig. 2** The individual order category scores by clinic, where each colored bar represents a clinic.

results for a patient, the linkage between abnormal lab values and the triggering of alerts seemed to be absent. As for the lack of drug monitoring alerts, physicians at these clinics commented that implementing these types of alerts at their clinics would be very helpful. Further research about the impact and use of these types of alerts in the outpatient setting could be beneficial for clinics so that they can implement them effectively into their EHR systems.

In terms of fatal order performance, the mean overall score was 67.9%, and all the clinics alerted on at least half of the fatal orders in their test. The clinics that detected the most fatal orders were also the clinics with the highest overall scores. This pattern is observed in the inpatient test, where there is a linear

relationship between overall performance in the test and fatal order performance.¹⁴ One of the fatal orders that clinics struggled with was the drug dose (daily) fatal order, where three clinics did not alert on it. Two of the clinics had daily dosing alerts implemented in their system (Clinics D and G), while Clinic B did not implement cumulative dosing alerts. For Clinics D and G, these results suggest that although they have cumulative dosing decision support implemented, it may not be targeting some of the most dangerous medication orders. Given the potential for these orders to cause significant harm and death, this is an important opportunity for improvement.

For nuisance orders, the mean overall score was 64.3%. These orders have the potential to contribute to alert fatigue

in that they are low priority drug–drug interactions or therapeutic duplications that can be presented noninterruptively.²¹ Overall, most clinics did not alert on many of these orders, however clinics with the highest overall scores had the lowest nuisance order scores, indicating that several nuisance orders were alerted on. For example, Clinic F had the highest overall score in the pilot (80%) but alerted on all the nuisance orders, suggesting that the threshold for which alerts are triggered is too low.²⁷ This relationship between overall score and nuisance order score was observed in the inpatient tool, where some hospitals achieved their high scores at the expense of overalerting.¹⁴ Other instances of overalerting in these clinics' EHRs included: repeated alerts for the same order and food and alcohol interactions being displayed within the same window as other potentially higher priority alerts. This can have significant effects on patient safety in that providers can miss severe alerts if they are delivered in the same window as low priority alerts.

Finally, three clinics had an EHR-based medication reconciliation functionality, while the other four clinics did not, even though medication reconciliation is required by the Joint Commission. However, during the pilot, only one of the clinics with the electronic medication reconciliation function could demonstrate this process. One clinic noted that though their system has this functionality, it is poorly understood by most providers, suggesting that either redesign or training sessions could be beneficial.²⁸ This process is very difficult if clinics cannot obtain the patient's most recent medication lists or do not have access to Continuity of Care Documents (CCDs). In this scenario, clinics without electronic medication reconciliation relied on information found in the last clinical note or on the patient bringing their medication list to their visit. Conversely, EHR-based medication reconciliation relies on consistent and accurate updates to the patient's medication list within the EHR,²⁹ and can also have consequences such as data entry errors.³⁰ For the clinics in the pilot, if patients were seen outside the health care system, their medication lists could not be updated automatically. This barrier suggests that the lack of interoperability of EHRs play significant roles in this process. Lastly, although medication reconciliation is useful for detecting discrepancies between patients' medication lists, it has yet to be proven to decrease the rate of ADEs.^{31,32} Clinics which had electronic medication reconciliation noted that when prescribers are updating the medication lists, no CDS is provided.

The results from this pilot provide insight into potential areas of improvement in outpatient EHRs. The clinics in the pilot have both basic and advanced decision support features implemented; however, some alerts related to advance decision support were completely turned off. The qualitative results from the pilot support the current research around medication reconciliation in that it is a useful process for detecting medication discrepancies, but the effectiveness of widespread implementation of electronic medication reconciliation on preventing ADEs has yet to be proven. These results suggest that there is significant variation in ambulatory EHR safety performance, and that ambulatory care provider organizations should engage in regular self-assessment to deter-

mine how well their EHR decision support system is preventing potential ADEs. Results from the inpatient tool shows that hospitals which take the test annually perform better than hospitals taking the test for the first time, suggesting that recurring evaluation can lead to improvement in EHR safety performance.⁷ Ultimately, this tool will be further developed to test other areas such as the usability of EHRs and alert design.

Next Steps for the Ambulatory EHR Evaluation Tool

For the next pilot, a software platform will be used to deliver all sections of test, and clinics will receive immediate feedback from each section of the tool. Next, at least 30 outpatient clinics will participate in this pilot, including the seven clinics that participated in this first pilot. There will be a change to the scoring algorithm, where nuisance orders will be included in the calculation of the overall score. This change is in response to the potential of nuisance orders to cause alert fatigue, which is a factor that can contribute to physician burnout.⁴

Limitations

Our study has several limitations. First, since EHR implementation varies by facility, the results from this tool are not representative of all the clinics using a specific vendor. In addition, this was the first pilot of the medication safety test, so it has not been externally validated; however, its methodology mirrors the inpatient test, which has been heavily validated.^{7,8,14–17} Next, this tool measures process quality rather than outcome quality and assessing the association between ADEs and scores on the pilot is outside the scope of this study. However, research on the inpatient tool has found an association between CDS performance and ADEs, where investigators predicted four fewer preventable ADEs per 100 admissions for every 5% increase in the overall percentage score.¹⁷ Fatal orders and nuisance orders belonged to certain order categories and are not representative of all the alerts a clinic might encounter. Within-vendor variation is common in hospitals, but this may not be true for outpatient clinics where there is less customization, and future studies should evaluate this. Finally, the clinics in our sample were recommended by EHR vendors and are likely to be more technologically advanced than the average clinic.

Conclusion

The Ambulatory EHR Evaluation Tool was piloted with seven clinics, and the results suggest that there are many opportunities for improvement. Clinics struggled in areas of advanced decision support such as geriatric alerts and none of the clinics alerted for the drug lab and drug monitoring orders. Given the potential for fatal orders to cause serious harm and death, clinics should seek to target these types of alerts first when seeking to improve their medication-related decision support and should strive for a perfect score in this area. At the same time, there must be a balance between underalerting and overalerting to help avoid alert fatigue. In the case of medication reconciliation, most clinics did not have an EHR-based medication reconciliation function and its accuracy is largely

dependent on accessing a patient's most updated medication lists. These data suggest that wider use of this tool by outpatient clinics could help improve an important dimension of medication safety and help inform vendors about areas of improvement as they work with their client bases.

Clinical Relevance Statement

Results from this seven-clinic pilot highlight some of the gaps in the implementation of ambulatory EHR systems, which can have serious effects medication safety. By using the tool, outpatient clinics can use their results to discuss improvements to their EHR with their vendor to prevent patient harm.

Multiple Choice Questions

- Based on this pilot, which of the following order checking categories did clinics struggle the most with?
 - Drug allergy.
 - Drug pregnancy.
 - Drug laboratory.
 - Drug–drug interaction.

Correct Answer: The correct answer is option c. None of the clinics in the pilot alerted for the drug laboratory orders in the medication safety test.

- Which order category did clinics have the most variation in performance in?
 - Drug laboratory.
 - Drug age.
 - Drug monitoring.
 - Drug allergy.

Correct answer: The correct answer is option b. Some of the clinics had drug age alerts turned on in their system, while others did not.

Protection of Human and Animal Subjects

No real patients were used in Ambulatory EHR Evaluation Tool, only test patients were used.

Funding

This study was funded by the Gordon and Betty Moore Foundation.

Conflict of Interest

D.W.B. consults for EarlySense, which makes patient safety monitoring systems. He receives cash compensation from CDI (Negev), Ltd, which is a not-for-profit incubator for health IT startups. He receives equity from ValeraHealth which makes software to help patients with chronic diseases. He receives equity from Clew which makes software to support clinical decision-making in intensive care. He receives equity from MDClone which takes clinical data and produces deidentified versions of it. He receives equity from AESOP which makes software to reduce medication error rates. He will be receiving

research funding from IBM Watson Health. Dr. Bates' financial interests have been reviewed by Brigham and Women's Hospital and Mass General Brigham in accordance with their institutional policies. All other authors have no competing interests to declare.

References

- Blumenthal D. Launching HITECH. *N Engl J Med* 2010;362(05):382–385
- 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(01):29–40
- Bates DW, Teich JM, Lee J, et al. The impact of computerized physician order entry on medication error prevention. *J Am Med Inform Assoc* 1999;6(04):313–321
- Gregory ME, Russo E, Singh H. Electronic health record alert-related workload as a predictor of burnout in primary care providers. *Appl Clin Inform* 2017;8(03):686–697
- Radley DC, Wasserman MR, Olsho LE, Shoemaker SJ, Spranca MD, Bradshaw B. Reduction in medication errors in hospitals due to adoption of computerized provider order entry systems. *J Am Med Inform Assoc* 2013;20(03):470–476
- Bates DW, Leape LL, Cullen DJ, et al. Effect of computerized physician order entry and a team intervention on prevention of serious medication errors. *JAMA* 1998;280(15):1311–1316
- Holmgren AJ, Co Z, Newmark L, Danforth M, Classen D, Bates D. Assessing the safety of electronic health records: a national longitudinal study of medication-related decision support. *BMJ Qual Saf* 2020;29(01):52–59
- Classen DC, Holmgren AJ, Co Z, et al. National trends in the safety performance of electronic health record systems from 2009 to 2018. *JAMA Netw Open* 2020;3(05):e205547
- Denham CR, Classen DC, Swenson SJ, Henderson MJ, Zeltner T, Bates DW. Safe use of electronic health records and health information technology systems: trust but verify. *J Patient Saf* 2013;9(04):177–189
- Kaushal R, Kern LM, Barrón Y, Quaresimo J, Abramson EL. Electronic prescribing improves medication safety in community-based office practices. *J Gen Intern Med* 2010;25(06):530–536
- Gandhi TK, Weingart SN, Seger AC, et al. Outpatient prescribing errors and the impact of computerized prescribing. *J Gen Intern Med* 2005;20(09):837–841
- Singh H, Spitzmueller C, Petersen NJ, Sawhney MK, Sittig DF. Information overload and missed test results in EHR-based settings. *JAMA Intern Med* 2013;173(08):702–704
- Ancker JS, Edwards A, Nosal S, Hauser D, Mauer E, Kaushal R with the HITEC Investigators. Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. *BMC Med Inform Decis Mak* 2017;17(01):36
- Co Z, Holmgren AJ, Classen DC, et al. The tradeoffs between safety and alert fatigue: data from a national evaluation of hospital medication-related clinical decision support. *J Am Med Inform Assoc* 2020;27(08):1252–1258
- Kilbridge PM, Welebob EM, Classen DC. Development of the Leapfrog methodology for evaluating hospital implemented inpatient computerized physician order entry systems. *Qual Saf Health Care* 2006;15(02):81–84
- Metzger J, Welebob E, Bates DW, Lipsitz S, Classen DC. Mixed results in the safety performance of computerized physician order entry. *Health Aff (Millwood)* 2010;29(04):655–663
- Leung AA, Keohane C, Lipsitz S, et al. Relationship between medication event rates and the Leapfrog computerized physician order entry evaluation tool. *J Am Med Inform Assoc* 2013;20(e1):e85–e90
- National Quality Forum. Safe Practices for Better Healthcare-2010 Update: A Consensus Report. 2010

- 19 The Leapfrog Group. Leapfrog Hospital Survey: Questions & Reporting Periods Endnotes Measure Specifications FAQs. 2020
- 20 Gandhi TK, Weingart SN, Borus J, et al. Adverse drug events in ambulatory care. *N Engl J Med* 2003;348(16):1556–1564
- 21 Phansalkar S, van der Sijs H, Tucker AD, et al. Drug-drug interactions that should be non-interruptive in order to reduce alert fatigue in electronic health records. *J Am Med Inform Assoc* 2013;20(03):489–493
- 22 Office of the National Coordinator for Health Information Technology. Non-federal Acute Care Hospital Electronic Health Record Adoption. Published 2017. Accessed January 11, 2019 at: <https://dashboard.healthit.gov/quickstats/pages/FIG-Hospital-EHR-Adoption.php>
- 23 Ratwani R, Fairbanks T, Savage E, et al. Mind the Gap. A systematic review to identify usability and safety challenges and practices during electronic health record implementation. *Appl Clin Inform* 2016;7(04):1069–1087
- 24 Dhillon-Chattha P, McCorkle R, Borycki E. An evidence-based tool for safe configuration of electronic health records: the eSafety checklist. *Appl Clin Inform* 2018;9(04):817–830
- 25 Wahls TL, Cram PM. The frequency of missed test results and associated treatment delays in a highly computerized health system. *BMC Fam Pract* 2007;8(01):32
- 26 Hysong SJ, Sawhney MK, Wilson L, et al. Understanding the management of electronic test result notifications in the outpatient setting. *BMC Med Inform Decis Mak* 2011;11:22. Doi: 10.1186/1472-6947-11-22
- 27 Weingart SN, Toth M, Sands DZ, Aronson MD, Davis RB, Phillips RS. Physicians' decisions to override computerized drug alerts in primary care. *Arch Intern Med* 2003;163(21):2625–2631
- 28 Scott K, Hathaway E, Sharp K, Smailes P. The development and evaluation of an electronic health record efficiency workshop for providers. *Appl Clin Inform* 2020;11(02):336–341
- 29 Wakefield M. Patient safety and quality: an evidence-based handbook for nurses. The quality chasm series: implications for nursing. In: Agency for Healthcare Research and Quality A. ed. Agency for Healthcare Research and Quality, AHRQ. Agency for Healthcare Research and Quality (US); 2008:1–1403. Accessed October 28, 2020 at: <https://www.ncbi.nlm.nih.gov/books/NBK2648/>
- 30 Wagner MM, Hogan WR. The accuracy of medication data in an outpatient electronic medical record. *J Am Med Inform Assoc* 1996;3(03):234–244
- 31 Tamblyn R, Abrahamowicz M, Buckeridge DL, et al. Effect of an electronic medication reconciliation intervention on adverse drug events: a cluster randomized trial. *JAMA Netw Open* 2019;2(09):e1910756
- 32 Mekonnen AB, Abebe TB, McLachlan AJ, Brien JAE. Impact of electronic medication reconciliation interventions on medication discrepancies at hospital transitions: a systematic review and meta-analysis. *BMC Med Inform Decis Mak* 2016;16(01):112