Mixed Results In The Safety Performance Of Computerized Physician Order Entry

Health Affairs 29, no.4 (2010):655-663
doi: 10.1377/hlthaff.2010.0160

The online version of this article, along with updated information and services, is available at:
http://content.healthaffairs.org/content/29/4/655

For Reprints, Links & Permissions : http://content.healthaffairs.org/1340_reprints.php
Email Alertings : http://content.healthaffairs.org/subscriptions/etoc.dtl
To Subscribe : https://fulfillment.healthaffairs.org
Mixed Results In The Safety Performance Of Computerized Physician Order Entry

ABSTRACT Computerized physician order entry is a required feature for hospitals seeking to demonstrate meaningful use of electronic medical record systems and qualify for federal financial incentives. A national sample of sixty-two hospitals voluntarily used a simulation tool designed to assess how well safety decision support worked when applied to medication orders in computerized order entry. The simulation detected only 53 percent of the medication orders that would have resulted in fatalities and 10–82 percent of the test orders that would have caused serious adverse drug events. It is important to ascertain whether actual implementations of computerized physician order entry are achieving goals such as improved patient safety.

Many people have suggested that electronic health records represent essential infrastructure for the provision of safe health care in the United States. For several years, the Institute of Medicine, the Leapfrog Group, the National Quality Forum, and other national groups concerned about patient safety have recommended, in particular, widespread adoption of electronic health records with computerized physician order entry.1–4

Background On Decision-Support Tools
With computerized physician order entry, physicians and other licensed clinicians write their orders for hospitalized patients electronically. This recommendation is based in large part on demonstrations by pioneering organizations. The organizations have shown that important improvements in safety can be achieved when rules-based decision support aids in averting medication errors and adverse events by providing advice and warnings as physicians write orders using a specially programmed computer.5–7

In this application of clinical decision support, physicians are made aware of potential safety issues that can result—for example, when ampicillin is given to a patient with a known allergy to penicillin, or the dose being ordered for a pediatric patient is much higher than the therapeutic range for a child of this age and weight. Prescribing errors such as these can lead to anaphylaxis or seizures, which are known as adverse drug events, if the medications are actually administered. The goal of medication safety decision support in computerized physician order entry is to prevent these types of serious errors as the orders are being written.

A study demonstrated that one in ten patients hospitalized in Massachusetts suffered an adverse drug event that could be prevented by decision-support tools in computerized physician order entry.8 The study spurred the passage of legislation requiring Massachusetts hospitals to implement computerized physician order entry by 2012 as a condition of licensure.9

More recently, “meaningful use” of computerized physician order entry and clinical decision support has been singled out as a requirement for hospitals to qualify for new financial incentives. These incentives will be offered under the...
Focus On Quality

health information technology (IT) stimulus provisions of the American Recovery and Reinvestment Act (ARRA) of 2009.

Computerized physician order entry interacts with other applications in the suite of digital tools that constitute the inpatient electronic health record (for example, to obtain information on allergies and patients’ weight) and is typically one of the later modules to be implemented. Hospitals’ adoption of the computerized physician order entry module is increasing, but slowly. Several reports have suggested that the successful application of decision support achieved among pioneering organizations is not being replicated and that implementation can create new problems. This report summarizes results for sixty-two hospitals across the United States that used a new simulation tool to assess their use of medication safety decision support in electronic health records with computerized physician order entry.

Study Data And Methods

HISTORY OF THE ASSESSMENT TOOL The impetus for developing the assessment tool was initially the standard developed by the Leapfrog Group. This is an employer group that seeks to accomplish breakthroughs, or “big leaps,” in hospital patient safety through a combination of public awareness and rewards to higher-quality providers. The group selected computerized physician order entry as one of the first three leaps in 2001. There was accumulating evidence concerning the frequency, tragic consequences, and financial costs of adverse drug events in hospitalized patients—and computerized physician order entry and decision support had demonstrated the ability to help avert many of them.

The Leapfrog standard includes two elements of meaningful use to ensure that computerized physician order entry has been implemented in such a way as to improve medication safety. According to the standards, physicians and other licensed providers must enter at least 75 percent of medication orders using computerized entry. Clinical decision support must also be able to avert at least 50 percent of “common, serious prescribing errors.”

Clinical decision support in this setting is the logic built into the computerized physician order entry system that, for example, checks to see if ampicillin has been ordered for a patient who is known to be allergic to penicillin. This tool was developed to specifically measure the ability of implemented electronic health record systems with computerized physician order entry to detect and avert these common yet serious prescribing errors in “live” hospital settings.

The tool was intentionally designed to give individual hospitals detailed and specific feedback on their performance and to give purchasers, through Leapfrog, an overall score for the hospital that can be used for benchmarking purposes.

This assessment complements efforts by the Certification Commission for Health Information Technology (CCHIT) to evaluate the capabilities available in vendors’ electronic medical record products “on the shelf.” It evaluates how products were implemented and are actually being used in hospitals.

The development of the tool was initially funded by the Robert Wood Johnson Foundation, the California HealthCare Foundation, and the Agency for Healthcare Research and Quality, and was completed in 2006. In April 2008 the assessment was incorporated into the Leapfrog Annual Safe Practices Survey for the first time, and hospitals completing the assessment received a feedback report. Beginning in 2009, assessment results were also factored into determining the extent to which the computerized physician order entry implementation met the Leapfrog standard.

DESIGN OF THE ASSESSMENT TOOL The assessment methodology is modeled after tools that are commonly used in other industries. It mimics what happens when a physician writes an order for an actual patient in the implemented electronic health record with computerized physician order entry. But it uses test patients—in effect, fictitious patients created for purposes of the assessment—and test orders.

A group of experts on adverse drug events, as well as the use of decision support in computerized physician order entry to decrease adverse drug events, developed test orders that are judged likely to cause serious harm (rather than those with low potential for harm). The test orders belong to the categories of adverse drug events (such as drug-to-allergy or drug-to-diagnosis contraindication) that prior research shows cause the most harm to patients. In most cases, they are actual orders that have caused adverse drug events, taken from primary adverse drug event data collection studies. The assessment offers a one-time, cross-sectional look at whether decision support provides advice to a physician writing such an order.

Decision support for this purpose is a set of tools or logic that can be integrated into the computerized physician order entry system to suggest appropriate orders (such as a dose calculator or a reminder to consider renal function) or to critique them once they have been entered, as through a message or an alert. The assessment
gives credit for all of these forms of decision support as relevant advice or information.

Because of differences in the epidemiology of preventable adverse drug events, different versions of the assessment were designed for adult and pediatric inpatient settings.

**USE OF THE ASSESSMENT TOOL** A designated team in the hospital performs a self-assessment in the following fashion. First, the team downloads instructions and information profiles for ten to twelve test patients. Then the team downloads around fifty test orders, instructions, and observation sheets to be used in the assessment.

A participating physician enters test orders for the test patients into the local electronic health record and observes and notes any guidance provided by decision support, such as the calculated dose, a message or alert displayed, and so on. The team enters the results obtained for each test order (decision support received or not). The assessment tool instantly computes an overall score (percentage of test orders identified), as well as the score for the orders in each adverse drug event category. It then displays results to the testing team. The entire process takes no more than six hours, and many hospitals complete it more quickly.

During the development period (2002–6), multiple testing was performed at more than twenty-five different hospitals. The testing reflecting all of the leading electronic health record and computerized physician order entry vendors’ products and ensured that the test could effectively evaluate each product. Details of the reliability and validity of the assessment methodology are available in the Online Appendix.

**Analysis**

Between April and August 2008, eighty-one U.S. hospitals completed the version of the assessment for adult patients. Test orders and posted results were reviewed, and information potentially identifying patients was removed. Nine hospitals were eliminated because registration information indicated that computerized physician order entry was being used only in the emergency department rather than more broadly in the hospital.

In addition, ten hospitals were excluded because they exceeded a deception-analysis threshold based on standard gaming detection strategies (testing irregularities such as exceeding time limits or multiple false positive results). A review of excluded hospitals did not reveal results that would have skewed the findings in this study. The resulting sample contained sixty-two hospitals.

The categories of adverse drug events addressed by the test orders include ones for which decision-support tools are fairly straightforward to implement (drug-to-drug or drug-to-allergy interactions, therapeutic duplication, inappropriate single dose, and inappropriate route of administration). The assessment also included test orders that require more effort to configure or customize—for example, use of an inappropriate daily dose or weight-based dose, drug-to-drug or drug-to-diagnosis contraindication, contraindications based on renal status or other metabolic abnormalities indicated by laboratory tests, or lack of monitoring.

The framework for basic and advanced decision support was based on prior work on clinical decision support in computerized physician order entry, and the categories are consistent with recent research on preventable adverse drug events.

**Statistical Analysis**

The basic unit for the analyses was the hospital. Overall scores for test orders and the scores for orders assigned to the “basic” group were both found to be approximately normally distributed. However, scores for test orders in the “advanced” group were slightly right-skewed. Thus, we present additional statistics to aid in interpretation. A fuller description of our analysis is available in the Online Appendix.

Tests based on standard assumptions about the population, such as through parametric tests, gave results almost identical to those of tests that made no such assumptions (nonparametric tests). Thus, for simplicity, parametric tests are displayed (see the Online Appendix for details).

The total hospital scores were found to be approximately normally distributed, as is typically the case with normal outcomes. Thus, we assumed a linear regression model with the overall score as the dependent variable (estimating the relationships between the outcome and covariates using ordinary least squares regression). Covariates that we identified as having a possible influence on total hospital score were vendor (there were nine), teaching status, hospital size groupings (number of beds), and whether or not the hospital was part of a health system.

We tested for appropriateness of our linear regression model using techniques that are described in the Online Appendix. We also used standard statistical techniques to determine the percentage of the overall variation explained by each factor.

We tested our model for appropriateness using techniques that are described in the Online Appendix. For dichotomous outcome variables,
such as test order detected (yes, no), percentages were calculated. However, to account for the hierarchical nature of the order data, or the test orders nested within hospitals, the 95 percent binomial confidence intervals for percentages were adjusted using the approach described by David Williams.23

All analyses were conducted using the statistical software package SAS 9.2. All tests were two-tailed, and a p value less than 0.05 (therefore not likely to be due to chance) was considered statistically significant.

Study Results

The types of hospitals included in the sample were broadly representative of larger U.S. hospitals (Exhibit 1). Nearly two-thirds were teaching hospitals. The higher representation of teaching hospitals and lower representation of smaller hospitals is consistent with the current pattern of adoption of computerized physician order entry in hospitals.10,11 Among the sixty-two hospitals, all but one reported using electronic health record applications including computerized physician order entry from one of seven commercial vendors.

**INDIVIDUAL HOSPITALS** Scores for individual hospitals ranged from 10 percent to 82 percent of test orders detected. The scores for the top 10 percent, or six hospitals, ranged from 71 percent to 82 percent. Scores for the six hospitals with the lowest scores ranged from 10 percent to 18 percent (Exhibit 2).

Mean hospital scores (Exhibit 3) were higher for orders that would lead to adverse drug events that can be addressed by basic decision support (61 percent) than for those requiring more advanced decision support (25 percent).

**AGGREGATE RESULTS: ADVERSE DRUG EVENTS** When results for all hospitals were pooled, the adverse drug event category detected most reliably was drug-to-allergy contraindication. Much higher scores were obtained for each of the categories addressed by basic clinical decision support than for those requiring advanced tools (Exhibit 4).

Drug-to-diagnosis contraindication includes pregnancy, which was also analyzed separately. These potential adverse drug events were only detected 15 percent of the time.

The set of test orders for each hospital includes four that are judged to result in patient fatality. When results for this subset were analyzed, we found that 47 percent (95 percent confidence interval: 36.9–57.6) were not detected by the decision support in use in these hospitals. Although hospitals do have pharmacy and nursing review processes in place that sometimes catch orders like these before the medication reaches the patient, these medication orders are far outside safe limits and would never be appropriate physician orders.

**CONTRIBUTING FACTORS** The information available for exploring contributing factors was limited to the vendor software solution in use, teaching status, hospital size by number of beds, and whether or not the hospital was part of a health system. We assessed the relationship between performance on the assessment and these factors.

High-low scores for hospitals using the same
computerized physician order entry software product ranged by as much as 40–65 percent (Exhibit 2). Some hospitals using each product detected at least 50 percent of the potential adverse drug events. The six top-performing hospitals used six different software products: one homegrown solution and five vendor products. In a multiple regression model, vendor choice was significantly correlated with performance \((p = 0.009, \text{ or not likely to be due to chance})\). This means that there is good statistical evidence to suggest that choice of vendors does have some positive effect on performance. However, vendor choice accounted for only 27 percent of the total.
Teaching status also correlated significantly with performance ($p = 0.007$, very unlikely to be due to chance), accounting for 10 percent of the observed variation in performance. But hospital size and being part of a hospital system did not correlate significantly with performance. At least as far as we can detect statistically, hospital size and being part of a hospital system do not influence performance one way or the other.

We tested for interactions between vendor and all other variables, and none were significant ($p > 0.2$). Finally, although we were able to account for only 39 percent of the observed variation in performance (overall $R^2 = 0.39$), standard tests (goodness-of-fit statistics and assessments) indicated that our model was appropriate for our study, as detailed in the Online Appendix.

### Discussion

Many of the benefits from inpatient electronic health records and the computerized physician order entry module in particular come from decision support. In this study, we found wide variation in the ability of implemented computerized physician order entry decision support to detect medication orders judged likely to cause serious harm to adult patients.

Many hospitals performed poorly, and the mean score was only 44 percent of potential adverse drug events detected. However, top-performing hospitals in this sample achieved scores of 70–80 percent or greater. To achieve these higher results, the hospitals have implemented advanced clinical decision support, as well as basic tools, and their performance provides a benchmark for the continuing efforts of lower-performing hospitals.

A comparison of aggregate scores with the findings from a recent study of adverse drug events, based on chart reviews in six community hospitals not using the computerized physician order entry module of the inpatient electronic health record, showed that decision support in the sixty-two hospitals is doing a much better job detecting adverse drug events that occur infrequently than those that occur more frequently.

The categories in our study with the three highest detection rates—drug-to-allergy, drug-to-drug, and duplicate medication—already together only contributed 7 percent of the adverse drug events in the community hospital chart review study. Fifty-five percent of the adverse drug events in the community hospital chart review study required considering laboratory results or patients’ age in determining medication appropriateness or dosing. For these categories in aggregate, corresponding test orders were detected only 20.6 percent of the time—not very good odds from the patients’ perspective.

### Different Products In Use

The high variability found, including among hospitals using the same electronic health record software product, emphasizes the importance of this type of assessment to gauge the extent to which decision support is being used. In fact, this is the rationale used by the Leapfrog Group and the National Quality Forum in including the use of the assessment tool in their recommended safe practices.
Hospital boards and executive teams who are investing in these systems, practicing physicians who adopt the technology in their routine work to improve patient safety, and external stakeholders requiring the use of electronic health records with computerized physician order entry or providing incentives for adoption all need a way to gauge progress in applying decision support to improve medication safety. This becomes acutely important as hospitals attempt to meet meaningful-use requirements to qualify for federal stimulus incentives.

**Assessing Patient Safety**

Given the large investment in achieving meaningful use nationwide, it will be important to ascertain whether actual implementation is achieving important goals such as improved patient safety. To our knowledge, this test is the first such objective evaluation of electronic health record systems in actual use.

The assessment measures the extent to which clinical decision support is providing some form of advice or an alert in response to medication orders that would cause an adverse drug event. The design and scope of decision-support capabilities in electronic health record software products vary, and local hospital configuration and customization are always required.

Although there was a relationship between the product involved and performance on the assessment, only 27 percent of the variation is associated with using different electronic health record products. This suggests that other factors had a bigger influence on the wide variation in performance among the hospitals. Although results are presented for one homegrown system, this is not a sufficient sample from which to draw any general conclusions.

In most hospitals, the use of decision support grows over time. At the outset, managing it is a new process. Including too much, poorly designed clinical decision support initially can even contribute to lack of computerized physician order entry adoption.

**Exploring Variability in Use**

Several factors are contributing to the variability observed. These include the completeness and ease of applying the decision-support tool sets in the electronic health record software; relevant knowledge and experience within the hospital; the availability and commitment of staff resources; and whether the use of clinical decision support began recently or has been honed over many years.

There are multiple possible explanations for the observed correlation between hospital teaching status and performance. These include such factors as research interest and having more staff resources to invest. However, these could not be explored further in this study.

For specific categories of adverse drug events, there are also multiple possible explanations for variability among hospitals. For drug-to-diagnosis contraindications, for example, many hospitals are likely not yet applying decision support because there is no constantly updated electronic problem list maintained by physicians for patients during their hospital stay.

The assessment also does not provide insight into whether or not the information or advice being presented is followed in practice—another prerequisite for improving medication safety. A better understanding of all of these factors is needed so that strategies for speeding progress can be devised.

Some reports have suggested one remedy: providing assistance to hospitals in using decision support effectively. This assessment tool provides guidance to the hospital in the form of a scorecard. Several publications exist to guide hospitals in their planning for and decisions about clinical decision support. Also, some vendors of computerized physician order entry software provide clinical decision support starter sets or facilitate sharing among customers.

Work is now under way to begin to develop libraries of decision-support rules and practical advice, although it is in early stages.

**Study Limitations**

This study has several limitations. Like all voluntary surveys, the study is subject to possible response bias. It is likely that the hospitals completing the assessment were more interested than many of their peers were in clinical decision support because it was initially offered as a learning exercise instead of being required for the annual Leapfrog survey.

Another limitation is that the sixty-two hospitals in the study represent about 8 percent of U.S. hospitals with the computerized physician order entry module of the inpatient electronic health record in use. Thus, they might not be representative. The fact that this is the first test of its kind means that there is no “gold standard”
for comparison of these results. However, during development of the assessment tool, the reliability and validity were extensively evaluated and found to be very high.

Finally, because results are self-reported, hospitals may have reported better performance than was actually assessed. However, it should be noted that public reporting of overall scores, which could influence hospitals’ behavior, was not implemented until 2009, and hospitals were asked to participate in part to aid them with improving their decision support. To address the potential for gaming, several standard safeguards were built into the test to detect patterns of use that suggest that gaming may be going on. Any assessment that included these characteristics was excluded from analysis.

Conclusions
In sixty-two hospitals using an inpatient electronic health record with computerized physician order entry, we found significant variability in the use of decision support to detect and provide advice or an alert concerning a medication order that would result in serious harm to an adult patient. Some hospitals performed very well, while others performed very poorly. In addition, the studied hospitals as a group were using basic decision support far more than the more advanced tools needed to detect types of orders that are major contributors to adverse drug events in chart-review studies.

These findings point to the importance of evaluations of the use of clinical decision support by hospitals to help guide their continuing efforts to improve medication safety. In addition, incentives to hospitals relating to computerized physician order entry should include some type of demonstration that clinical decision support is actually being employed, instead of being based solely on whether computerized physician order entry is in use.

The broader use of this type of assessment of meaningful electronic health record use should be explored for other software applications used in direct clinical care.

The authors thank Peter Kilbridge, Fran Turisco, and the project expert advisers, as well as individuals in many hospitals, health systems, and software vendor companies, for their involvement and assistance in some phase of the development and testing of the assessment tool used in the study. They also thank the Leapfrog Group for access to the data for analysis. The development of the assessment tool used in this study was supported by grants from the California Healthcare Foundation, the Robert Wood Johnson Foundation, and the Agency for Healthcare Research and Quality (Contract no. 290-04-0016, Subcontract no. 6275-FCG-01). These organizations did not sponsor and were not involved in the reported analysis or preparation of the manuscript. Four of the authors (Jane Metzger, Emily Welebob, David W. Bates, and David C. Classen) were involved in the development of the assessment tool. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agency.

NOTES

20 The Online Appendix can be accessed by clicking on the Online Appendix link in the box to the right of the article online.