

Analysis of a Failed Clinical Decision Support System for Management of Congestive Heart Failure

Rajiv Wadhwa MD, Douglas B. Fridsma MD PhD, Melissa I. Saul MS, Louis E. Penrod MD, Shyam Visweswaran MD PhD, Gregory F. Cooper MD PhD, Wendy Chapman PhD
Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh

Abstract

In order to increase compliance with The Joint Commission's Congestive Heart Failure Core Measures, a rule based clinical decision support system (CDSS) was developed and deployed at a community hospital in our health system. We evaluated the performance of the CDSS in identifying patients with primary congestive heart failure (CHF) and identified problems encountered with its introduction. Performance of the CDSS was compared against a manual review of records of patients with diagnosis of primary CHF. The CDSS had a sensitivity of 0.79 and PPV of 0.11. The CDSS issued multiple alerts for majority of the patients (74%). The number of alerts issued for patients without primary CHF was large, and for a majority of patients (63%) physicians did not respond to alerts the first time. The CDSS performed poorly and was eventually withdrawn but provided insight into a subsequently successful method for managing CHF.

1. Introduction

The quality of healthcare provided across the hospitals in US has been a cause for concern for some time now¹. The concerns include frequent medication errors^{2,3} delay in the introduction of clinical research findings into medical practice⁴, and poor adherence to preventive health recommendations. Deployment of electronic medical records (EMR's) with computer-based physician order entry (CPOE) capability, coupled with clinical decision support systems (CDSS) has been advocated as a possible solution to these problems⁵⁻⁷.

As a part of its overall quality improvement efforts, the Joint Commission has developed 36 evidence-based, clinically validated measures that are mandated for all hospitals in the country⁸. These measures are referred to as the Core Measures, and four of them apply to management of congestive heart failure (CHF). At the University of Pittsburgh Medical Center (UPMC), we piloted a CDSS application aimed at improving compliance with Core Measures for CHF management at one of our community hospitals with a fully integrated EMR.

The system failed. In this paper, we describe development and implementation of the application, an evaluation of the system's performance, and an error analysis for determining why the application did not succeed. We conclude by describing the lessons we learned from the failure and changes we have made that have ultimately resulted in increased compliance with the Core Measures.

2. Background

Before the introduction of the CDSS, compliance with CHF Core Measures at University of Pittsburgh Medical Center's (UPMC) affiliated hospitals, ranged between 60% and 90%. The process in place for identifying eligible patients relied on a dedicated nurse who manually reviewed patient records to confirm a diagnosis of CHF and appended paper reminders on charts to encourage physicians to complete the Core Measures, if they had not already been completed.

In early 2006, UPMC decided to implement a rule-based CDSS aimed at increasing compliance with the CHF Core Measures. It was believed that a CPOE-based CDSS would help (1) identify patients for whom the Core Measures applied, (2) encourage real-time documentation needed to comply, and (3) assist physicians in ordering necessary medications or tests. A multi-disciplinary decision support team was formed that developed rules and an interface for the CDSS. The CDSS was piloted in a single community hospital that had shown ready adoption of the EMR including CPOE. The results of the pilot were to be used to guide implementation of the CDSS in the entire hospital system.

The piloted application went live on August 14, 2006 and demonstrated problems immediately after implementation. The decision support team received enough complaints that the rule that triggered most alerts was deactivated three weeks after going live. The entire system was withdrawn after fewer than eight weeks in use.

After evaluating why the system failed, the decision support team decided to take an entirely different approach, making use of many of the rules they had

implemented but reconfiguring them into a different workflow. Instead of generating alerts to physicians, the rules are now used to classify patients into 'Possible' and 'No' CHF. Patients with 'Possible' CHF are included in a daily report that is sent to a case manager. The case manager examines the patient records and decides whether a patient is likely to be coded as CHF. For potentially positive patients, the case manager flags the paper records and contacts physicians if the Core Measures have not already been met, recommending necessary orders for complying with the measures. The new approach has been much very successful, boosting the compliance rates with Core Measures to above 95%.

In this paper we (1) describe how the CDSS was designed to work, (2) evaluate how well the system performed at identifying CHF patients, and (3) assess the problems encountered with the system to better understand why the CDSS failed.

3. Methods

In this section, we explain how the CDSS was designed to work by describing the computerized rules and the desired workflow of the CPOE system. We also describe the evaluation we performed to understand how the system worked and the problems that were encountered with the pilot implementation.

3.1 Core Measures and brief description of the CDSS

Four Core Measures apply to patients diagnosed with primary CHF:

1. Assess left ventricular function (LVF) for all patients with CHF.
2. For patients with impaired LVF, treat with angiotensin converting enzyme inhibitor (ACEI) or angiotensin receptor blocker (ARB) or document a written contraindication to both.
3. Complete discharge instructions for the patients.
4. Provide advice regarding smoking cessation.

The goal of the CDSS was to identify patients admitted to the hospital who were likely to be discharged with a diagnosis of CHF and then prompt physicians to complete the Core Measures for those patients. The CDSS development team intended to accomplish this in the following four steps:

1. Identify patients with CHF: The team developed five rules to help identify patients with CHF. The developers envisaged that the majority of CHF patients could be identified if they had a documented past history of CHF or if some of the commonly used medications or tests for CHF patients were noted

Table 1. Identifying Rules and their triggers.

History Rule	IF a past history of CHF was recorded in the Problem List or CHF was a diagnosis in a prior Discharge List
Chest X-ray Rule	IF a Chest X-ray was ordered within 28 days of an echocardiogram (ECHO)
ECHO Rule	IF an ECHO was ordered within 7 days of a Chest X-ray
IV Lasix Rule	IF a single dose of intravenous furosemide or bumetanide was ordered
AICD Rule	If an order for AICD was entered.

AICD: Automated implantable cardioverter defibrillator

during their hospital stay. Five identifying rules were implemented and are described in Table 1. All rules fired only for patients who were not on an ACEI or ARB agent.

2. Alert physicians to CHF patients. Once a CHF patient was identified by the firing of one or more rules in Table 1, the CDSS alerted the physician by generating a CHF Alert (Figure 1) when the patient's chart was opened.

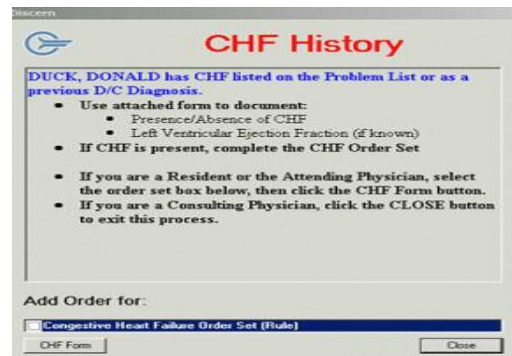


Figure 1. CHF Alert for the History Rule.

3. Encourage physicians to document CHF. On getting an alert the physician was expected to click on the 'CHF Form' button on the CHF Alert, which would then display the CHF Form (Figure 2). The physician would check 'Yes' on the CHF Form to confirm the presence of CHF or 'No' if CHF was absent. Once a CHF Form was completed no further CHF Alerts would be generated for that patient. However, if the CHF Form was not completed, the CDSS generated alerts on subsequent instances of the chart being opened.

4. Expect physician to complete the CHF Order Set. If a physician checked 'Yes' in the CHF Form, he or she was also expected to order an ACEI or ARB or document a contraindication to these agents. To do this, the physician was expected to return to the CHF

Alert and open the CHF Order Set from a pull-down menu (Figure 3), and complete it.

Figure 2. CHF Form

Component	Order Detail
Medications	
****Ace Inhibitor****	
<input type="checkbox"/> ramipril	5 mg, Cap.
<input type="checkbox"/> lisinopril	2.5 mg, Tab
<input type="checkbox"/> fosinopril	10 mg, Tab
<input type="checkbox"/> enalapril	2.5 mg, Tab
****Angiotensin Receptor Blocker****	
<input type="checkbox"/> losartan	50 mg, Tab
<input type="checkbox"/> valsartan	80 mg, Tab
<input type="checkbox"/> candesartan	4 mg, Tab.
****ACE/ARB meds are not ordered, complete order below with reason not ordered included****	
<input type="checkbox"/> Cardiac Med Not Ordered Reason	T.N, CHF, A
<input type="checkbox"/> Cardiac Med Not Ordered Reason	T.N, CHF, A

Figure3. CHF Order Set

3.2 Evaluating Performance of the system

To assess the performance of the CDSS we created a gold standard (GS) by manual record review. One author (RW) manually reviewed medical records of all patients discharged with a diagnosis of primary CHF during the period of time that the CDSS was live during the pilot phase (from 7/24/06 to 9/12/06). We first confirmed the diagnosis of CHF, and then examined the records and cardiac reports to see if a patient's LVF had been documented. For patients who had a documentation of LVF, we next determined if they had an impaired LVF. If so, we determined if they had been discharged on an ACEI/ARB agent or if a contraindication to both these agents had been documented.

We examined the log of rules that fired for all patients in the study period. Since the GS identified a list of patients for whom the rules should have fired, we were able to determine the following: True Positives (TP) – patients for whom rules fired and who had primary CHF; False Negatives (FN) - patients for whom rules did not fire but who had primary CHF; and False positives (FP) – patients for whom rules fired but who did not have primary CHF.

From the available numbers we calculated sensitivity and positive predictive Value (PPV).

3.3 Problems encountered with the CDSS

We performed an analysis of three types of problems with the CDSS to gain an understanding into what went wrong with the implementation.

1. False negative classifications. We examined the records of patients who had CHF but were not identified by the CDSS to assess why the rules did not identify positive patients.

2. Excessive alerts. Physicians complained that there were too many alerts being generated and interrupting their workflow. Since the History rule (Table 1) was turned off after three weeks because it appeared to generate most of the excessive alerts, we studied alerts for the time period 7/24/06 to 8/13/06 only. There were three sources for false positive rule firings: the initial firing for a negative patient, multiple subsequent firings for the same negative patient, and multiple subsequent firings for positive patients whose initial alert was a true positive. We measured the number of times a rule fired for a negative patient (FP), the number of primary CHF patients who had multiple alerts, and the total and average number of alerts per primary CHF patient

3. Incomplete physician response to alerts. For some patients the CHF Forms were completed in response to alerts but CHF Order Sets were not. We first identified patients for whom the CHF Form had been checked as 'Yes'. Next we studied these patients' records to find out if they had been discharged on an ACEI or ARB agent or a contraindication to both had been documented in their records.

4. Results

We evaluated performance of the CDSS and examined missed and excessive alerts.

4.1 Performance of the CDSS

There were a total of 112 patients identified in the GS with a discharge diagnosis of primary CHF. Rules fired for 54 and did not fire for 58 of them. Of these 58 patients, 44 were on an ACEI or ARB agent, which would have prevented any firings, leaving 14 false negatives. Therefore, the sensitivity of the CDSS was 0.79 (54/68). One or more of the rules fired for a total of 503 patients, of which 54 were true positives, resulting in a PPV of 0.11 (54/503). Table 2 shows the contingency table for patients for whom the rules did and did not fire.

Table 2. Contingency table for CDSS Performance.

	Pts. For whom rules should have fired	Pts. For whom rules should not have fired	Total
Pts. For whom rules fired	54 (TP)	449 (FP)	503
Pts. For whom rules did not fire	14 (FN)	unknown	
Total	68		

4.2. Problems encountered with the CDSS

1. False negatives. Identifying rules did not fire for 14 primary CHF patients who were not on an ACEI or ARB agent. For eight of the 14 patients, we found no triggers for any of the rules in their records, suggesting that these patients could not have been identified by the existing set of rules. For the remaining six of 14 patients we did find triggers for rule firings.

2. Excessive alerts. Excessive alerts include FP alerts (patients for whom rules fired but who did not have primary CHF) and multiple alerts for patients with primary CHF (which were appropriate). There were 3023 rule actions for 449 patients who did not have primary CHF. We did not have the frequency distribution of individual rules that for this group and were therefore unable to calculate the exact number of alerts for this group, although it was likely to be large. Multiple alerts were generated for seven of 16 primary CHF patients for the first three-week period. Overall, the rules fired an average of 3.6 times per patient, with the History rule firing 4.4 times per patient, and the other rules 1.2 times (Table 3).

Table 3. Alerts for primary CHF patients for the first three weeks.

Rule(s)	Number of alerts	Number of patients	Avg. alerts per patient
History Rule	53	12	4.4
Other rules	6	5	1.2
Total	59	16	3.6

3. Incomplete physician response to alerts.

Physicians checked ‘Yes’ for 16 primary CHF patients. However, six of the 16 patients (38%) were not discharged on an ACEI or ARB agent and were not documented with a contraindication, indicating that the CHF Order Sets were not completed for those patients.

5. Discussion

1. Problem of False negatives. Overall sensitivity of the CDSS (0.79) was moderate but still suboptimal. There is no single test, medication or clinical observation that is 100% sensitive for CHF. Rules were built on some of the commonly used orders for drugs and tests in the management of CHF but covered very specific instances of the tests or drugs used (for example, IV Furosemide x1 only). Sensitivity could have been increased by including more general orders for tests and drugs (for example, Furosemide 40 mg bid) but this would have decreased the PPV even further. So it is not surprising that a few CHF patients were not picked up by the rules. The main failing of the CDSS was due the large number of false positives leading to a low PPV of 0.11 and due to the generation of multiple alerts for a single patient.

2. Problem of Excessive Alerts. The false positives included not only patients who did not have CHF but also patients with a diagnosis of secondary CHF. The distinction between primary and secondary CHF is made only retrospectively by the coders. In many ways the characteristics of patients are similar in primary and secondary CHF patients. Therefore it is to be expected that the conditions for triggering of the rules would often be met in secondary CHF patients. Another explanation is that the rules as framed were rather ‘non-specific’ in that they would have triggered in a variety of clinical situations other than CHF. The IV Furosemide rule for example, would have fired with Furosemide x 1 order, which may be given in patients with dependent edema or hyponatremia.

The multiple alerts for the same patients occurred because the CHF Forms were not completed by physicians. If a physician did not complete the CHF Form in response to an alert and instead closed the alert, another alert could be generated in a subsequent logon. Repetitive alerts contributed to ‘alert fatigue’ which in turn may have lead physicians to ignore any further alerts, compounding the problem. There are several possible explanations for not completing the forms. Some physicians likely felt, depending on their specialty, that they were not responsible for managing the patient’s CHF and therefore chose to close the Form. For an overburdened physician on call, bypassing the CHF Form altogether might have been an attractive option. A more appropriate timing for alerts may have elicited a better response rate from these physicians. Moreover, completion of the CHF Form was a multi-step, cumbersome process that may have dissuaded some physicians from completing it, especially those that were less computer savvy.

3. Incomplete physician response to alerts.

Of 16 patients with completed CHF Forms, physicians completed Order Sets for only ten of them. Not completing Order Sets implied that these patients did not have their Second Core Measure met, defeating the entire purpose of the CDSS. The problem of incomplete forms was partially due to the assumption that after completing the CHF Form (Figure 2), physicians would next check and open the CHF Order Set and complete it. In many cases, physicians did not follow this procedure. Completion of CHF Form without completing the CHF Order Set suppressed any further alerts. As a result, many physicians learned that it was not necessary to complete the Order Set in order to suppress alerts. Perhaps it would have been better if the system would have allowed suppression of further alerts only after completion of the CHF Order Set and if the Order Set were to appear automatically instead of relying on the physicians to open it.

6. Conclusion

Overall, the pilot implementation of the CDSS to improve compliance with CHF core measures performed poorly, leading to its withdrawal. To identify all or most patients with CHF is a difficult task. In this trial the rules covered very specific instances of ordered drugs and tests. If the rules would have been more inclusive, covering more general orders for drugs and tests, they would have identified nearly all CHF patients. This would have caused a further increase in false positives, generating yet more alerts and compounding the problems encountered. Therefore, while a CDSS should have high sensitivity, it also needs to have a high PPV to be an effective tool.

A complementary approach would be to use statistical and machine learning techniques to learn better predictive models from variables identified by domain experts. However, learning of such models requires a large number of patient cases that have been labeled appropriately by domain experts. Using NLP techniques in identifying patients could be yet another approach that may increase PPV.

A better designed, more streamlined and user friendly interface for CHF Alert and Form may have increased the response rate^{9, 10}. A greater emphasis on physician education emphasizing the usefulness and necessity of the CDSS along with the demonstration of completion of the CHF Alerts, Forms and Order Sets could have helped improved the response rate

significantly¹¹. Ultimately, we learned that CDSS's may succeed better if they avoid alerts aimed at physicians and develop a methodology that fits better in the existing workflow.

Acknowledgments

This research was funded in part by a grant from the National Library of Medicine (5T15-LM007059).

References

1. McGlynn, E. A. et al. The quality of health care delivered to adults in the United States.[see comment]. *NEJM* **348**, 2635-45 (2003).
2. Bates, D. W. et al. Incidence of adverse drug events and potential adverse drug events. Implications for prevention. ADE Prevention Study Group.[see comment]. *JAMA* **274**, 29-34 (1995).
3. Kaushal, R. et al. Medication errors and adverse drug events in pediatric inpatients.[see comment]. *JAMA* **285**, 2114-20 (2001).
4. Antman, E. M., Lau, J., Kupelnick, B., Mosteller, F. & Chalmers, T. C. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction.[see comment]. *JAMA* **268**, 240-8 (1992).
5. Chertow, G. M. et al. Guided medication dosing for inpatients with renal insufficiency. *JAMA* **286**, 2839-44 (2001).
6. Dexter, P. R. et al. A computerized reminder system to increase the use of preventive care for hospitalized patients.[see comment]. *NEJM* **345**, 965-70 (2001).
7. Garg, A. X. et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review.[see comment]. *JAMA* **293**, 1223-38 (2005).
8. Commission, C. a. T. J. in *Version 2.3* (2007).
9. Chan W. Increasing the success of physician order entry through human factors engineering. *J Healthcare Inform Manage* **16**, 71-9 (2002).
10. Johnson, C. M., Johnson, T. R. & Zhang, J. A user-centered framework for redesigning health care interfaces. *Journal of Biomedical Informatics* **38**, 75-87 (2005).
11. Bates, D. W. et al. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. *JAMIA* **10**,523-30(2003).