PREDICTING THE IMPACT OF AN ELECTRONIC HEALTH RECORD ON PRACTICE PATTERNS USING COMPUTATIONAL MODELING AND SIMULATION

Thomas R. Clancy, Ph.D., RN, Connie White Delaney, Ph.D., RN, FAAN, FACMI, Alberto Segre, Ph.D., Kathleen Carley, Ph.D., Andrew Kuziak, Ph.D., Hwanjo Yu, Ph.D.

Abstract

The overall purpose of this research study is to discover and apply new knowledge regarding methods to predict the impact of an electronic health record (EHR) on clinical practice guidelines in complex systems such as hospitals. Specifically, the aims of this study are: 1) to build, simulate and validate the accuracy of a computational model representing the current practice patterns in a sample of patients diagnosed with heart failure (HF) and treated in a community hospital; and 2) using computational modeling and simulation, develop a method to predict the effects of best practice guidelines on practice patterns after implementation of an EHR.

The model results showed a significant decline (p < .05) in net tests and treatments of 26.22%, total annual process hours of 12.56% and costs by $121,513. Total nurse work-time declined by 30% while physician work time increased by 88%. The results of this study suggest that implementation of an EHR with embedded clinical guidelines can improve the utilization of resources and process delay times in a population of 245 heart failure patients.

Introduction

In an effort to reduce cost and improve quality, hospitals have been automating clinical and administrative processes by transitioning to electronic health records (EHR). The benefits of an integrated suite of computer applications that comprise an electronic health record (EHR) have been researched extensively1,2,3,4. Benefits include improved legibility, decreased medical errors, access to decision support and expert systems, multiple access points (hospital, physician office, mobile devices) and better adherence to best practice clinical guidelines.

Although promising, the adoption of EHRs has been slow. For example, it is estimated that only 9.6% of hospitals have an electronic health record with a fully functional computerized provider order entry system4. Barriers to adoption include excessive order entry and documentation time, increased workload, interference with workflow patterns, limited access, frequent updates, computer glitches, difficulty for people with underdeveloped keyboarding skills and cost5.

Given the external pressures to improve quality and reduce expenses but faced with the substantial cost of EHR implementation and the uncertainty of results, healthcare administrators find themselves caught in a dilemma. How can one predict the effects of information technology on such performance improvement strategies as the adoption of best practices clinical guidelines, medication error reduction, implementation of standardized nursing languages, and use of expert systems for decision support before having to purchase them? In this study a new application, computational modeling and simulation, is introduced as a methodology for predicting the impact of an EHR on a sample of 245 patients admitted to a community hospital with a diagnosis of heart failure.

Methods

The following research question was addressed in this study; Is there a significant difference in the actual, total annual number and process times (hours) of laboratory tests, radiology procedures, medications administered, days of care provided, respiratory care services, ECG’s and echocardiograms for a sample of 245 heart failure patients admitted to a community hospital after implementation of an EHR with embedded guidelines?

Procedures

Modeling Current State Variables: Data from 285 patients coded with a primary diagnosis of heart failure were selected to build as the “training group” for the current practice computational model. This sample consisted of the entire population of heart failure patients treated as acute inpatients for the calendar year 2003.

AMIA 2007 Symposium Proceedings Page - 145
Using secondary data, the total number of lab tests, radiology procedures, medications administered, direct nursing care hours delivered, days of hospital stay, respiratory care, electrocardiograms, echocardiograms and mortality were entered for each patient on an Excel spreadsheet. These tests were chosen because they represented 80% or greater of total tests and treatments provided during the patient’s hospital stay.

The model was built using Arena, a commercial, discrete event simulation application. Heart failure patients entering the model arrived at a constant inter-arrival delay of 1.28 days between admissions which provided a total of 285 HF discharges annually. All HF patients entered a nursing unit admission assessment “Process” module with delay times drawn from an IID (independent and identically distributed) triangular probability distribution function (PDF).

After admission to the nursing unit, the model used a “ReadWrite” module to sequentially import (from an Excel spreadsheet) the actual quantity of physician orders by department for each HF patient’s total length of stay. The models simulated outputs were validated against the actual data by comparing the total number of lab tests, radiology procedures, medications ordered, patient days, electrocardiograms (ECG) and echocardiograms (echo), and respiratory care services in one year between the empirical values and simulated output. There was not a significant difference between means of the actual values and the simulated output (five replications) at p < .05.

To simulate process times, each patient’s total quantity of orders flowed to the appropriate department (lab, radiology, pharmacy, respiratory therapy ECG and echos) where they were multiplied by per order processing time. Per order processing times were drawn from each department’s actual average process delay time per test or treatment for that year. The total number of direct care nursing hours per HF patient was determined from GRASP, a commercially available software application used for predicting staffing needs. GRASP utilizes standardized times for common nursing tasks that have been validated through time and motion studies.

Physician charting time was determined from time and motion studies of eleven physicians in seven different specialties during morning rounds. Labor costs associated with each department were determined by dividing the total annual salary expenses by total annual staff hours for each department.

Modeling Heart Failure Guidelines: To integrate best practices into the model, the American College of Cardiology and the American Heart Associations Guidelines for the Evaluation and Management of Chronic Heart Failure in the Adult \cite{6} was referenced. For each individual patient in the sample the overall percent compliance with the guidelines was calculated. Compliance with a guideline was accepted if the recommended test or treatment was ordered at least once during the patient’s length of stay.

To identify those physicians with the overall best practice (the BP group), the sample was sorted by severity index\textsuperscript{7}, patient volume, compliance with guidelines, mortality and length of stay. Each performance outcome was rank-ordered within its own category and then a composite score was calculated for each physician. The top 25% of desirable scores (highest severity index, highest patient volume and highest compliance but with lowest morality and length of stay) was clustered within three individual physicians (65 patients). It was noted that the BP group had a significantly higher severity index, lower mortality rate and lower length of stay (p>.05) when compared to the remainder of the sample.

In constructing the best practices computational model the control logic required physicians to either follow practice guidelines or consult with the BP group when ordering tests and treatments. Specifically, patients acted as inputs to the model and were automatically fed case by case from the Excel spreadsheet (through a Visual Basic interface in Arena) over a simulated year. As each patient entered the model it flowed into a “Decision Module” that acted as a gatekeeper. The module (using Boolean operators) compared the actual, total number of tests or treatments ordered for each patient (lab, radiology, pharmacy, nursing hours, patient days, ECGs, echos and respiratory care) against the median value for the same variable in the BP group. If the total number of tests or treatments was less than or equal to the median, the value would pass through the module and be recorded. If the number exceeded the median, the model would draw from a theoretical probability density curve fitted to that variables entire distribution in the BP group.

The model represented embedded clinical practice guideline order sets that if exceeded would prompt a consultation from the best practices physician group. If the BP group was consulted, the number of tests and treatments recorded were then drawn from the BP group’s probability density curve. The final
Training the Model: Once the best practice model was complete it was necessary to “train” it so that the simulated utilization of tests and treatments did not vary significantly from the actual values in the BP group. A total of 197 patients (70% of the original sample) were selected as the training group while the remaining 88 patients (30%) were set aside to be used later as a test group. Training the model consisted of iteratively adjusting thresholds in each variable’s Decision Module to produce a simulated annual output that was not significantly different than the actual data in the BP group. To validate the training model against the BP group the 197 patients were run over a simulated year (times 5 replications) and the means between similar variables in each group were compared using a Student’s t-Test.

Verifying the Model: To test the model using different patients but from the original sample, the severity index between the “hold-out” group of 88 patients and the training group were first compared using a Students “t” Test. Results indicated there was no significant difference in severity between the two groups at the p<.05 level. The holdout group of 88 patients was run over a simulated year (times 5 replications) and the means between the same variables in BP group were compared using a Student’s t-Test. To validate the model using patients from a different population, a new sample of 245 patients with a primary diagnosis of heart failure were selected from the hospital’s database. The severity index between the training group and the new population was compared using a Student “t” Test. There was no significant difference in severity between the two groups at the p<.05 level. The 245 patients from the new population were then run through the model over a simulated year (times 5 replications) and means between similar variables in the BP group were compared.

Running the Virtual Experiment: To initiate the virtual experiment, the previous test sample of 245 heart failure patients were run through the model over a simulated year (times 5 replications) with all parameters set to the current pre-EHR state (with no clinical guidelines). The parameters were then set to the post-EHR state (which included clinical guidelines) and the same 245 patients were run through the model over a simulated year (times 5 replications). The means of all variables (total annual lab tests, radiology procedures, medications administered, patient days, NICVS tests, respiratory care services as well as process delay times) pre and post EHR implementation were then compared using paired t-confidence intervals.

Results

The average overall compliance with guidelines pre EHR was 58.01%. The results of the virtual experiment are presented in Table 1. The simulated output data indicated there was a significant decrease in the predicted total annual number of lab tests, radiology tests, medications administered, days of care, and respiratory care services after implementation of an EHR with embedded guidelines for the 245 patients. There was not a significant drop in ECG or echo tests post EHR implementation. The area with the greatest drop post EHR implementation was in medications administered which dropped from 29,100 to 15,729 annually. The overall average weighted decline in tests and treatments was 26.22%.

The predicted change in total annual process delay times (hours) post EHR implementation is presented in Table 2. There was a significant drop in total annual hours for lab, radiology, pharmacy, respiratory care and nurse work time. However, there was a significant increase in total annual physician work time, increasing from 207.05 to 388.75 hours. There was not a significant difference in delay times for ECGs and echos. Total annual nurse work time showed the greatest overall decline dropping from 8,374.28 hours to 5,892.01 hours. The net average weighted impact on process times for all variables was a decline of 12.56%.

The impact on physician and nurse work time post EHR implementation without embedded guidelines was also evaluated. Table 3 presents the total annual change in physician and nurse work time post EHR implementation without embedded clinical guidelines. By removing the drop in tests and treatments attributable to guidelines, the model predicted a significantly greater increase in physician work time but less of a decline in nurse work time.

The Arena software also provides a costing module that computes total labor costs. The total annual cost of labor to process all tests and treatments pre-EHR implementation was $523,753. After implementation of the EHR with guidelines, total labor costs dropped to $402,240 for a predicted total annual savings of $121,513.
Discussion
The study represents an idealized model which prompts physicians to either follow heart failure guidelines or consult with a specialist if certain thresholds are exceeded (such as length of stay). This may accurately represent the impact of protocol driven guidelines. In fact this model was used to aid in determining the impact of a hospitalist program (which utilized protocol driven guidelines) on overall resource consumption in the hospital.

Although there was a significant decrease in the number of tests ordered and process cycle times, much of the gains in utilization and productivity were simply the result of a reduction in length of stay. This suggests that the primary value of an EHR is in enabling the adherence to best practice guidelines. The model also highlighted the ongoing dilemma between local and global optimization. Physician workload increased as a result of using an EHR but overall cycle time, resource consumption and cost decreased.

The value in using a computational model to predict the impact of an EHR lies in its ability to run accurate but inexpensive scenarios. Adherence to guidelines and consultations to heart failure specialists could be altered in the model depending on the local environment. Thus a hospital can customize the model to better fit its current practice patterns.

Limitations
The model used a number of assumptions that had a bearing on the accuracy of the simulated outputs. For instance, determining compliance with the AHA/ACC guidelines for management of heart disease from the available data was problematic. The small sample sizes of the training, holdout and test group may have also had an impact on the accuracy of the model.

Conclusions
The results of this study suggest that implementation of an EHR with embedded clinical guidelines can improve the utilization of resources and process delay times in a population of heart failure patients. Future research using the model in this study should focus on validating the predicted model outputs with actual data collected post EHR implementation. Research should also concentrate on validating the model on patient populations with diagnoses other than heart failure. The model should provide the opportunity to test various workflow scenarios that would improve physician work time after EHR implementation.

References
Table 1
Comparison of Total Annual Tests and Treatments Pre and Post EHR Implementation with Embedded Guidelines (*significant at p < .05)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Pre-EHR (actual)</th>
<th>Post-EHR (simulated)</th>
<th>Difference</th>
<th>+/- 95% CI for Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lab Tests</td>
<td>3,735</td>
<td>3,237 (498)*</td>
<td>47.83</td>
<td>47.83</td>
</tr>
<tr>
<td>Radiology Tests</td>
<td>462</td>
<td>353 (109)*</td>
<td>30.98</td>
<td>30.98</td>
</tr>
<tr>
<td>Medications Administered</td>
<td>29,100</td>
<td>15,729 (13,371)*</td>
<td>1,005</td>
<td></td>
</tr>
<tr>
<td>Days of Care</td>
<td>1,272</td>
<td>979 (293)*</td>
<td>53.04</td>
<td></td>
</tr>
<tr>
<td>ECGs &amp; Echos</td>
<td>538</td>
<td>512 (26)</td>
<td>21.21</td>
<td></td>
</tr>
<tr>
<td>Respiratory Care Services</td>
<td>3,748</td>
<td>2,534 (1,214)*</td>
<td>39.09</td>
<td></td>
</tr>
</tbody>
</table>

Table 2
Comparison of Total Annual Process Delay Time by Area, Pre and Post EHR Implementation with Embedded Guidelines (*significant at p < .05).

<table>
<thead>
<tr>
<th>Variables; Process Delay (hours)</th>
<th>Pre-EHR (simulated)</th>
<th>Post-EHR (simulated)</th>
<th>Difference</th>
<th>+/- 95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD Work Time</td>
<td>207.05</td>
<td>388.747</td>
<td>181.69*</td>
<td>1.62; 18.77</td>
</tr>
<tr>
<td>Nurse Work Time</td>
<td>8,374.28</td>
<td>5,892.01 (2,488.27)*</td>
<td>183.6; 263.3</td>
<td></td>
</tr>
<tr>
<td>Lab</td>
<td>1,195.2</td>
<td>1,040.02 (155.18)*</td>
<td>6.82; 15.17</td>
<td></td>
</tr>
<tr>
<td>Radiology</td>
<td>716.1</td>
<td>585.91 (130.19)*</td>
<td>0; 50.24</td>
<td></td>
</tr>
<tr>
<td>Pharmacy</td>
<td>1,164</td>
<td>632.85 (529.15)*</td>
<td>0; 40.11</td>
<td></td>
</tr>
<tr>
<td>Respiratory Care</td>
<td>862.08</td>
<td>589.18 (272.91)*</td>
<td>33.71; 39.09</td>
<td></td>
</tr>
<tr>
<td>ECGs and Echos</td>
<td>694.020</td>
<td>678.65 (15.37)</td>
<td>0; 26.3</td>
<td></td>
</tr>
</tbody>
</table>

Table 3
Comparison of Total Annual Physician and Nurse Work Time Pre and Post EHR Implementation (No embedded guidelines). *Significant at p < .05.

<table>
<thead>
<tr>
<th>Variables; Process Delay (hours)</th>
<th>Pre-EHR (simulated)</th>
<th>Post-EHR (simulated)</th>
<th>Difference</th>
<th>+/- 95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD Work Time</td>
<td>207.05</td>
<td>489.47</td>
<td>282.42*</td>
<td>1.62; 4.093</td>
</tr>
<tr>
<td>Nurse Work Time</td>
<td>8,374.28</td>
<td>7,226.802</td>
<td>1,147.48*</td>
<td>183.6; 146.1</td>
</tr>
</tbody>
</table>