Design and Evaluation of a Web-Based Interactive Visualization System for Lung Transplant Home Monitoring Data

David S. Pieczkiewicz, MA¹, Stanley M. Finkelstein, PhD¹, Marshall I. Hertz, MD²
¹Division of Health Informatics, Department of Laboratory Medicine and Pathology
²Division of Allergy, Pulmonary, and Critical Care Medicine, Department of Medicine
University of Minnesota, Minneapolis, MN

Abstract
We describe the implementation and evaluation of an interactive system for displaying patient pulmonary information from a lung transplant home monitoring program. Software was designed to take case information and render it as Scalable Vector Graphics (SVG) in a web browser. Twelve clinician readers reviewed twenty randomly-ordered, retrospective cases in each of three display formats (graphical, graphical interactive, and tabular) and determined whether they showed signs of infection or graft rejection. Decision times and reader preferences were also recorded. No statistically significant differences were found with respect to decision accuracy. However, the graphical displays yielded up to 25% faster decision times than numerical tables. Readers overwhelmingly preferred the graphical formats, particularly the interactive displays. We believe that graphical, interactive displays of patient data would be well-accepted and efficacious tools in clinical practice, whether for transplant care, or any care involving the assessment of large bodies of time-oriented, multivariable data.

Introduction
Graphical means of communicating quantitative and qualitative information have been used in health care for decades. Technological advances have allowed for the increasing sophistication of these visualizations, including the development of interactivity, the ability of users to manipulate visualizations to obtain additional information. Visualization and human-computer interaction are active areas of research, and guidelines have been published to encapsulate “best practices” in data visualization, using principles of human perception and cognition¹⁴. Published studies suggest that the formats used to display information can have an impact on the quality and timeliness of medical decisions, with the specific, optimal formats depending on the nature of the data and decision tasks⁵⁻⁷. With few exceptions⁶, existing studies have focused on simplified decision tasks, often with synthetic data and removed from clinical contexts. The research presented here, part of a larger study of medical data visualization, investigates the performance of several types of data presentation in a specific clinical decision context: that of determining the infection/rejection status of lung transplant recipients.

Lung transplant recipients are placed on lifetime regimens of immunosuppressive medications to prevent graft rejection, which can also leave them susceptible to respiratory infections⁹,¹⁰. Traditionally, infection and rejection are detected through clinic visits, where patients perform spirometric maneuvers to record lung function. More recently, home monitoring technologies have been used to allow for more timely detection of clinical problems. The Lung Transplant Home Monitoring Program (LTHMP) at the University of Minnesota has supplied participating patients with portable electronic spirometers for home use. Each day, pulmonary function and symptom data are recorded, and are uploaded each week to caregivers who monitor the data for incipient infection or rejection. If problems are noted, the patient’s primary care provider is notified, and the appropriate actions are taken. This home monitoring system has been found to be well-accepted, efficacious, and cost-effective¹¹⁻¹³, and has become a local standard of care.

The time-series data examined by clinicians in this triage task are typically presented in the form of numerical tables, and form the motivation for the present study. For such data, detecting trends of decreasing lung function is the primary decision task, and research has suggested that line graphs may be a better presentation format in this case, as they group time series data into lines which can be more easily apprehended perceptually and cognitively as single objects, rather than the many individual objects of numerical tables¹⁴. Further, interactive displays which allow users to obtain more information from the data may also allow for more accurate decision making. For this study, we developed several different display formats (modalities) and compared their performance using clinician readers to investigate the hypothesis.
would allow for more accurate and rapid detection of infection and rejection events than would numerical tables.

Methods and Materials

The experiment was implemented as a series of cases displayed to clinician readers, who gave judgments as to each case’s infection/rejection status. Cases were 90-day series of demographic, pulmonary, and symptom information taken retrospectively from the LTHMP databases. Case status (positive or negative for infection/rejection) was determined by independent clinician review of the full medical record of each sampled case. Ten positive and ten negative cases were used, for a total number of cases \( C = 20 \). Twelve clinician readers (\( R = 12 \)) participated in the study. The readers were pulmonary physician fellows, transplant physicians, nurse transplant specialists, and pulmonary technicians, with an average clinical experience of 17.8 years.

Three display modalities (\( M = 3 \)) were designed and implemented: a numerical table, a “static” (non-interactive) line graph, and an interactive line graph. All three modalities presented the same underlying data fields, based on consultations with domain experts: demographic and clinical information, including age, sex, transplant type (e.g., single-lung, double-lung, etc.) and underlying condition (e.g., COPD, CHF, etc.); pulmonary function information, including forced expiratory volume at one second (FEV\(_1\)), forced vital capacity (FVC), forced expiratory mid-flow rate (FEF\(_{25-75}\%\)) and peak expiratory flow rate (PEFR); and subjective, self-reported symptom information, including temperature, shortness of breath at rest, wheezing, coughing, and sputum amount and color. Symptom information was reported on 5-point Likert scales (e.g., 1 = no wheezing, 5 = severe wheezing, etc.).

The interactive graphical modality was based on the static graphical modality, and included the ability to mark and query specific points for further information, obtain comparative information between two points, draw a straight line between two points, and obtain a five-day moving average as an overlay on a graph.

All three formats were implemented using a PHP script to take database data and generate the displays in Scalable Vector Graphics (SVG), an Extensible Markup Language (XML) application for representing vector graphics information in a textual format. The SVG was then rendered in a web browser using the Adobe SVG plug-in. The interactive effects and calculations in the interactive graphical modality were implemented as client-side JavaScript, which manipulated the Document Object Model (DOM) of the SVG and its web page. A partial view of the tabular modality is shown as Figure 1, and a full view of the interactive graphical modality with some interactive features in use is shown as Figure 2.

The experimental design was balanced, in that each case was viewed by each of the clinician readers in each of the modalities (i.e., each reader viewed 3 modalities \( \times 20 \) cases = 60 displays). All displays were presented to each reader in random order, and shown in single viewing sessions. Decisions were recorded as probabilities on a scale of 0-100, with 0 representing a definite negative case (no infection or rejection present), and 100 a definite positive. Decision times were recorded with a stopwatch. At the end of each viewing session, readers were asked to rank each of the displays in terms of subjective preference.

The decision probabilities were used to obtain receiver operating characteristic (ROC) curves for each modality, which were then analyzed using multiple reader-multiple case analysis of variance (MRMC ANOVA), a technique commonly used in radiology studies comparing two or more imaging modalities. Decision times were analyzed via mixed-effects modeling of the log-transformed time measures (corrected for right-skewing), with reader and case as random effects, and modality as a fixed effect.

<table>
<thead>
<tr>
<th>Lastname, Firstname</th>
<th>MID: XXX - TABLE</th>
<th>Single-lung Tx for IPF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Transplanted XX XXX XXX</td>
</tr>
<tr>
<td></td>
<td></td>
<td>XX weeks, X days since transplant (as of 22 Aug 2000)</td>
</tr>
</tbody>
</table>

Figure 1. Partial view of tabular modality, slightly reduced. The full modality extends to 90 days.
Figure 2. Interactive graph modality, slightly reduced, showing several interactive features in use, including “brushing” of data points by the cursor (i.e., “tooltips”), marking of points with horizontal lines, comparison of two points (most recent two marked points compared to the right of each graph), virtual straightedge (a line drawn between the most recent two marked points), and five-day moving average. Apart from the interactive features and associated buttons, the static graph modality is identical.
Results

Accuracy and time measures, along with 95% confidence intervals, are presented as Table 1.

<table>
<thead>
<tr>
<th>Modality</th>
<th>AUC</th>
<th>95% CI</th>
<th>Time (sec)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interactive Graph</td>
<td>0.65</td>
<td>[0.46, 0.84]</td>
<td>45.3</td>
<td>[33.5, 61.2]</td>
</tr>
<tr>
<td>Static Graph</td>
<td>0.67</td>
<td>[0.48, 0.86]</td>
<td>39.7</td>
<td>[29.3, 53.6]</td>
</tr>
<tr>
<td>Table</td>
<td>0.66</td>
<td>[0.47, 0.84]</td>
<td>53.6</td>
<td>[39.7, 72.5]</td>
</tr>
</tbody>
</table>

Table 1. Summary accuracy and time measures.

While the static graph modality has the highest overall area-under-the-curve (AUC) measures, which can be read as the probability of a correct positive/negative classification, the confidence intervals around these accuracy measures are wide, and the MRMC ANOVA did not reveal a statistically significant difference between the three modalities ($F_{2,22} = 0.147; P = 0.86$). Differences in decision times, however, were more apparent. The interactive graph modality was approximately eight seconds (15%) faster than the tabular modality, and the static graph was approximately 14 seconds (26%) faster than the table. The overall F-test for decision times was highly statistically significant ($F_{2,687} = 24.45; P < 0.0001$), and multiple comparisons revealed that all three modalities were statistically significantly different from one another.

Eleven out of the twelve readers ranked the interactive graph modality first in terms of preference (mean rank = 1.1), and nine ranked the table modality last (mean rank = 2.75). Eight of the readers ranked the static graph second (mean rank = 2.2).

Discussion

No statistically or clinically significant differences were found in terms of decision accuracy between the three modalities. The reasons for this, and for the low overall accuracies, are still under investigation, but may be related in part to the ambiguity of the clinical decision task, which requires readers to identify any of several different types of graft dysfunction, which may have their own distinct signs and symptoms. Also, different readers may be applying different mental models to the decision task, and to the relative importance of different data elements. The lack of established clinical guidelines for detecting infection and rejection events may also play a role. Finally, as the readers were not given the full case information that was used in formulating the gold standard, the low accuracies may reflect a fundamental difficulty of detecting rejection and/or infection on the basis of the clinical parameters provided.

The differences in decision times were significant, however, and lend support to the hypothesis that graphical means of presenting time-series data may decrease decision times, by organizing the information in a manner which is more conducive to trend detection. The slightly longer decision times for the interactive graph modality, compared to the static graph modality, may be due to readers spending extra time using the interactive features to explore the data and confirm emerging clinical judgments.

Reader preferences were most striking, and showed an overwhelming preference for the interactive graph modality. Readers made frequent use of the interactive features, when available, particularly the ability to draw horizontal lines through points, and obtain comparative information between two points. In their comments, readers said they liked the quick overviews of case data afforded by the graphical modalities. A common request was for additional clinical information in the timelines, such as dates of hospitalization or changes to the patient’s medication regimen.

The visualization system developed here can display case information in any web browser with SVG support, either through plug-ins or direct rendering, making the system largely platform-independent. While the scripts used to build the displays read in data from simple, tab-delimited text files prepared ahead of time, they could easily be changed to dynamically query and extract data directly from databases, using PHP’s software hooks to many database management systems. Due to these implementation advantages, as well as the successful results of our experiments, the system is currently being adapted for day-to-day clinical use at the Fairview-University Transplant Center.

Conclusion

Technological advances are allowing for the collection and storage of ever-increasing amounts of multivariable medical data, such as in electronic medical records. New methods of visualization are needed to help manage this abundance of information. Based on theories in information visualization and cognitive psychology, we designed and implemented several ways of displaying clinical data from a lung transplant home monitoring system.
program. Results indicated that while there were no statistically significant differences in terms of decision accuracy between the modalities tested, decision times using the graphical modalities were significantly faster. Clinician readers showed a clear preference for the graphical modalities, particularly the interactive graph. We conclude that graphical, interactive presentations of pulmonary and other clinical information are efficacious, rapid, and well-accepted means for the determination of infection and rejection events in lung transplant recipients.

Acknowledgements

The research described here represents part of the primary author’s dissertation research, supported in part by the Biomedical Engineering Institute at the University of Minnesota. The Lung Transplant Home Monitoring Program was funded by NIH grant RO1-NR02128. We gratefully acknowledge the assistance of John Troiani, MD, who reviewed the cases used. Most of all, we thank the clinician readers for their time and suggestions, and the patients who participated in the LTHMP.

References