Innovative Information Visualization of Electronic Health Record Data: a Systematic Review

Vivian West, David Borland, W. Ed Hammond

February 5, 2015

Duke Center for Health Informatics
Outline

• Background
• Objective
• Methods & Criteria
• Analysis & Findings
• Limitations
• Conclusions

Background

- 2004 Presidential executive order: ‘Electronic Health Records (EHR) for all Americans
  - Provide accessible EHR for most Americans within 10 years
- 2009 Health Information Technology for Economic and Clinical Health Act (HITECH Act)
  - Allocated $19.2 billion in incentives to increase use of EHRs
- Centers for Medicare and Medicaid Services (CMS) incentive payments from CMS for adopting EHRs
History of Data Visualization

- William Playfair
  - First to use basic graphical visualization of data
  - Stated readers best understand and retain information by graphical representations of data
  - 1786: The statistical breviary: shewing, on a principle entirely new, the resources of every state and kingdom in Europe; illustrated with stained copper plate charts, representing the physical powers of each distinct nation with ease and perspicuity

Napoleon’s March on Russia in 1812
Published by Charles Minard in 1861

Florence Nightingale, 1858

The areas of the blue, red, & black wedges are each measured from the centre as the common vertex. The blue wedges measured from the centre of the circle represent areas for which the deaths from Preventable or Mitigable Zymotic diseases; the red wedges measured from the centre: the deaths from wounds; & the black wedges measured from the centre: the deaths from all other causes. The black line across the red triangle in Nov. 1854 marks the boundary of the deaths from all other causes during the month. In October 1854 & April 1855, the black area coincides with the red; in January & February 1855 the blue coincides with the black. The entire areas may be compared by following the blue, the red, & the black lines enclosing them.


Death by lack of sanitation
Death by wounds
Death by Other Causes

Large outer gray bands represent deaths due to lack of sanitation
Lighter gray middle/inner bands represent death from wounds during the war
Darkest middle/inner bands represent death by other causes
Visualization Techniques

• Graphs for vital signs
• Fishbone diagrams: laboratory results
• Knowledge discovery using
  – Large and small data sets
  – Scales, shapes, colors
  – Bar charts, line graphs, scatter plots, pie charts
• “Information visualization”-- Interactive visual representations of abstract data to amplify cognition
• Finance, accounting, and the petroleum industry: account for volume and complexity of data
• Visualization techniques to large and complex EHR datasets limited
Health Care Data Visualizations in the 1990’s

- Summarize patient status
- Use several diverse data sets in EHR to visualize information
- Plots of test results & treatment data (Pownser & Tufte, 1994)
- Clinical records contain longitudinal data of patient visits over time with records of changing problems, medications, treatments, and responses related to health status
- Graphs illustrate data so comparisons, trends, and associations can be understood
- Healthcare studies use graphs with time as the horizontal axis
  - Visualization tools developed using temporal data
LifeLines (Plaisant et al.\textsuperscript{4})

- Began with user interfaces for Juvenile Justice Information Systems
- Graphical attributes
  - Colors and lines depicting a patient’s discrete events

LifeLines/LifeLines2
Plaisant, et. al.

Visualizing Patient Records

Discovering Temporal Categorical Patterns Across Multiple Records


Knowledge-based Navigation of Abstractions for Visualization and Explanation
Visualizing Health Care Data

• Longitudinal data from EHRs displayed through innovative visualization techniques has tremendous potential for discovering useful information in data.
• Before EHR, little emphasis on using large and complex datasets.
• LifeLines/LifeLines2 & KNAVE/KNAVE-II/VISITORS (Visualization of Time-Oriented Records) most widely reported.
• Longitudinal studies in PubMed have increased from 7,071 publications (1983) to 45,821 (2013).
• EHR data is a new kind of data that requires new visualization techniques to discover knowledge.
Objectives

• Investigate visualization techniques that have been used with EHR data and answer the following questions:
  – What is the prevalence of the use of information visualization with EHR data?
  – Are techniques being used for knowledge discovery with an entire EHR dataset?
  – What has been learned from research on visualization of EHR data?
Methods

• Conducted a systematic literature review following PRISMA (Preferred Reporting Items for Systematic reviews and Meta-analysis)

• Limited articles between 1996 and 2013
  – Veterans Health Administration first mandated use of EHRs
  – Health Insurance Privacy and Portability Act (HIPPA) enacted
  – First study using visualization with complex data published by Plaisant, et. al.
Methods

• Electronic literature search conducted in May – July 2013 using MEDLINE & Web of Knowledge
• Supplemented by citation searching and gray literature searching
• Used broad keywords to assure comprehensive document search

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Boolean</th>
<th>Additional keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information visualization</td>
<td></td>
<td>Health data, electronic health record, electronic medical record</td>
</tr>
<tr>
<td>Information visualization</td>
<td>AND</td>
<td></td>
</tr>
<tr>
<td>Visualization</td>
<td>AND</td>
<td>Big data, clinical data, health data, health care data, healthcare data, electronic health record, electronic medical record</td>
</tr>
</tbody>
</table>
Inclusion Criteria

• Use **EHR data** and **innovative visualization techniques** or describe techniques applicable to EHR data
• Includes articles describing static and interactive visualizations of EHR data
• Definitions:
  – **EHR data** – electronic clinical records containing clinical information (eg, demographics, problems treatments, procedures, medications, labs, images, providers)
  – **Innovative visualizations** – visualizations other than standard graphs; must use complex and large data
Exclusion Criteria

• Articles are excluded if:
  – Related to animals or plants
  – Did not describe specific techniques for visualization
  – Solely described the need for visualizations
  – Focused on non-EHR topics:
    • Technical details related to visualization
    • Genetics
    • Syndromic surveillance
    • Geospatial environmentally aware data

Article Selection and Analysis

Medline (PubMed & PMC) and Web of Knowledge

Records identified through database searching (n = 847)

Additional records identified through other sources (n = 44)

Duplicates removed (n = 191)

Records screened (n = 700)

Records excluded (n = 666)

Full-text articles assessed for eligibility (n = 34)

Did not meet Inclusion Criteria

Full-text articles excluded
- With reasons (n=16)
  - NA (9)
  - Viz NA (1)
  - GIS (1)
  - Genetics (1)
  - Position Paper (3)
  - Technical (1)

Studies included in qualitative synthesis (n = 18)
Article Selection and Analysis

- Collected in Excel
- Read first 50 abstracts and titles
- 11 themes identified
- Created matrix
- Read 34 → 18
- No statistical analyses

Objective: investigate prevalence of info regarding techniques used for EHR data
## Analysis

<table>
<thead>
<tr>
<th>No.</th>
<th>Visualization Discussed</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>LifeLines\textsuperscript{11, 14-16}</td>
<td>Most advanced application; provides timeline of a patient’s temporal events</td>
</tr>
<tr>
<td>1</td>
<td>LifeLines2 \textsuperscript{19}</td>
<td>Enables use of multiple patient records; users see both numerical and categorical data; evolves into LifeFlow \textsuperscript{31}, used for millions of patient records to understand trends</td>
</tr>
<tr>
<td>4</td>
<td>VISITORS \textsuperscript{17-20} (evolved from KNAVE/KNAVEII)\textsuperscript{12,18}</td>
<td>Evolved from KNAVE/KNAVEII to accommodate diverse temporal data from multiple records; system feasible for exploring longitudinal data but interface needs simplification</td>
</tr>
<tr>
<td>1</td>
<td>Radial Starburst \textsuperscript{34}</td>
<td>Shows complexity of data represented over 100-dimensional space; has potential for interactive system</td>
</tr>
</tbody>
</table>
## Analysis

<table>
<thead>
<tr>
<th>No.</th>
<th>Visualization Discussed</th>
<th>Usability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pattern Matching &amp; Temporal Interpolation(^\text{29})</td>
<td>Might have been excluded; used EHR data for 3 million patients; showed value in using time in correlation &amp; aggregated data from many records vs a single record</td>
</tr>
<tr>
<td>1</td>
<td>DICON (Dynamic Icon)(^\text{30})</td>
<td>Interactively explore clusters of similar patients; clusters are represented as icons on a treemap; unique, but requires time for users to understand</td>
</tr>
<tr>
<td>1</td>
<td>Outflow(^\text{32,33})</td>
<td>Looks at disease progression paths; allows users to look at a visual display consisting of multiple events, their sequences, and outcomes</td>
</tr>
<tr>
<td>5</td>
<td>Other: ie. Treemaps,(^\text{28,30}) radial displays,(^\text{34,36,38}) icicle trees.(^\text{31})</td>
<td></td>
</tr>
</tbody>
</table>
Analysis (n=18)

- 16 focus on use of visualizations for clinical decision support;
- Other 2 suggest use for quality assurance and improvement
- 15 studies address the use of temporal data
- Most describe interactive visualizations
- Reported training time ranged from 6 to 30 minutes
Common Themes & Challenges

• Themes:
  – Type of data accessible to the user
  – Meaningfulness of visualizing large amounts of data
  – Usability
  – Training time

• Challenges:
  – Data (quality, size, diversity)
  – Users (needs, skills)
  – Design (ability to visually explore & analyze results)
  – Technology (tools, infrastructure)
Challenges

• The amount of EHR data and its display is a challenge
  – Difficult to see and identify meaningful patterns in visualizations
  – Zoom, pan, and filter tools reduce clutter but will not suffice for ‘big data’

• Size & complexity of EHR data is a challenge
  – Color, density, and filtering techniques distinguish variables
  – No reported techniques discuss applicability to entire datasets from EHR and potential for knowledge discovery

• Ability to use temporal data in visualizing aggregate data from EHRs is important to users

• Need design that presents a single interactive screen
Challenges

• Awareness of many variables that can lead to uncertain data in EHRs, potentially distorting temporal events
• Complications from missing values, inaccurate data, and mixed data types
• Users want to see categorical and numerical data both on a large scale and in detail
• Normalization scheme for aggregated numerical data
• Training time considerations, complexity of the data
Limitations of Study

- Review limited to primary publications describing innovative visualization techniques and application to EHRs
- Intentionally broad search terms
- Eliminated articles with more technical abstracts and a focus on geospatial representation
- No books
Conclusions

• Few techniques effectively and efficiently display large and complex data in EHRs
• Need techniques to handle Big Data & temporal data
• Look to techniques from other disciplines
• Research to date has identified important findings that can help guide future research
Funding

Supported by the US Army Medical Research and Material Command (USAMRMC) grant number W81XWH-13-1-0061
Thank you!