Overview: Visualizing Healthcare Data – A Literature Survey

Julia Zhu
Background

• Vast quantities of digital healthcare data are generated
  • Clinical trials, patient information, research findings, etc.

• Complexity in processing information increases
  • Leads to information overload

• Healthcare industry beginning to use visualization tools to visualize data
Overview

• My paper aims to:
  • Identify past studies that have used visualization idioms for health data, with an emphasis on multivariate data
  • Discuss how the idioms were used in these studies
  • Their advantages and disadvantages

• Health data visualization is still a fairly new subject

• 23 papers were found
  • Majority were design studies
  • Spanned from 2007 – 2016
  • Mostly published after 2013
Visualizing Multivariate Health Data

• Multivariate data = data involving 3 or more variables

• Visualization idioms:
  • Radar plot
  • Tree map
  • Bubble chart
Radar Plots

• Used in clinical studies to show changes over time

• Structure:
  • Circular graph with centrally-projecting rays
  • Each spoke represents a different variable
  • Ray lengths represent the values of the variables

• Saary’s 2007 study examined how radar plots could be used to visualize patient satisfaction with their interaction with the health care system

Figure 1. Example image of a radar plot.
Saary’s Radar Plot

- Scale from dissatisfied (1) to totally satisfied (5)
- Total satisfaction is represented by the perimeter (outermost ring).
- The radar plot shows that patients from Town B are generally more satisfied than that other two.

Figure 2. Radar plot comparing 3 groups on 7 variables.
Treemaps

- Used to visualize genome data or health data statistics

- **Structure:**
  - A window subdivided into multiple parts
  - Each part represents unique data
  - Area of rectangle proportional to a specified dimension of the data

- Hugine et al.'s 2014 study researched how surgeons' surgery performances in comparison to their counterparts could be visualized with treemaps

*Figure 3. Example image of a treemap.*
Hugine et al.’s Treemap

- Surgeon ID on the upper left corner of each block
- Size of the rectangles represents the length of stay, in days, for each patient
- Color of each rectangle represents the outcome for each patient
  - Red = deceased
  - Green = alive

Figure 4. Treemap of surgeries performed by surgeons, and the surgery outcomes for each patient.

Bubble Charts

- Used to compare the relationships between data objects in 3 numeric-data dimensions

- **Structure:**
  - Data points represented as bubbles
  - Data represented on x axis, y axis, by bubble size and/or bubble color

- Al-Hajj *et al.*’s 2013 study used a bubble chart to visualize data and to spot trends on hospitalization from injury

*Figure 5. Example image of a bubble chart.*
Al-Hajj et al.’s Bubble Chart

- Y axis: cost of hospitalization
- X axis: hospitalization rate
- Bubble size: injury cost
- Bubble color: type of injury

Figure 6. Bubble chart visualization of hospitalization data showing hospitalization rates and hospitalization costs per injury cause.

Multidimensional Health Data & Geospatial Regions

• Population healthcare data are usually tightly paired with geospatial regions

• Visualization idioms:
  • Spiral theme plot
  • Ring Map
Spiral Theme Plot

• Used to visualize many years of patient data

• **Structure:**
  • A spiral pattern composed of stacked categories (themes)
  • Themes are along a spiral curve - the time axis
  • Each data point plotted within themes

• Bloomquist *et al.* 2016 visualized seasonal flu patterns of patients over 4 years with a spiral theme plot

*Figure 7. Example image of a spiral theme plot.*
Bloomquist et al.’s Spiral Theme Plot

- 1 dot = 1 patient
- Dot radius = patient’s age
- Dot color = patient’s race
- Dot shape = patient’s gender
- Time = spiral base curve
- The theme’s width at a particular angle = total occurrence of flu at that time

Figure 8. Spiral theme plot displaying seasonal pattern of flu over four years.

Ring Maps

• Used to visualize multivariate spatial health data
• Illustrates individual datasets as separate rings of information

• Structure:
  • Concentric, segmented rings -- circular or elliptical
  • Rings surround a base map of a certain geographic region of interest

• Lopez-De Fede et al.’s 2011 study used a ring map to visualize county-level HIV/AIDS, and STD diagnostic data for South Carolina

Figure 9. Example image of a ring map.
Lopez-De Fede et al.’s Ring Map

• Ring map: central base map and ring display
• Central base map shows South Carolina’s 46 counties
• 1 spoke = 1 county
• Ring display has 4 concentric rings
• 1 ring = a separate layer of data
• Ring shade = illness prevalence

Figure 10. Ring map showing diagnosis rate of HIV/AIDS, syphilis, gonorrhea, and chlamydia in South Carolina counties.

Advantages & Disadvantages

• The strengthens and limitations for each type of idiom is different
• In general they can be categorized as:
  • **Advantages:**
    • Effectively present the bigger picture in a single diagram
    • Display many variables without using 3D graphs
    • Can accommodate longer time studies
  • **Disadvantages:**
    • Can exhibit high information density
    • Layout design can affect decision-making
    • May be difficult to interpret
Thank you!