HEALTH NETWORKS AND DISPARITIES

Primary Investigator: Jim Moody, PhD
Project Manager: Laura Sheble, PhD
Molly Chen, Emily Wu

Medicine tends to be problem-centric, focusing on conditions as separate. Is there a way to improve the comprehensiveness of diagnoses through a holistic analysis: which diagnoses tend to co-occur in patients of similar demographic groups?

**AIM:** Construct an interactive diagnosis network that allows users to explore patterns between disease categories and demographics as well as narrow down which health conditions tend to co-occur

**DATASET:**
1,048,576 entries (2007-2011)
Duke hospital records, Durham Cty.
Grouped by demographics:
- Age: 0-90+
- Race: White, Black, Asian, Latino
- Gender: M | F
Diagnoses coded using ICD-9 codes
*In order to construct node-edge pairs, all patients kept in dataset were diagnosed with at least 2 conditions*
CONSTRUCTING THE NETWORKS

Global Network

- Parse CSV → JSON format, create node (diagnosis) and edge (patient) lists with source-target pairs and indices
- Run data in Gephi to create network with layout, color codes
- Add Sigma JS plugin to make network interactive

Ego Network

- Consist of focal node and neighbors
- Use force-directed D3 layout to map nodes near related disorders
- Uses search algorithms to match user input with corresponding diagnosis entry
- Displays neighboring diagnoses

Health Networks and Disparities

- Each node represents a diagnosis
- Each edge represents a patient
- Earlier diagnoses → Later diagnoses

User enters search query

MySQL request pulls query from database

Server passes info along to network/user
User Search Algorithms

1. **String Match** – match user input word-by-word; least flexible since user needs to know medical vocabulary
2. **Boolean Search** – combine components of words and return intersection; returned too many results
3. **Google Custom Search** – load search results from customized search: [www.icd9codes.com](http://www.icd9codes.com); working solution, accounts for context of search/similar terms

**ICD-9 Categorization**

- Allows users to easily determine which disease category their diagnosis falls into
- We added to the existing ICD-9 hierarchy, which organizes 3-digit codes into more specific 5-digit codes.
- We came up with categories → subcategories → diagnosis.