Using Medicare Claims to Study Patient-Sharing Networks of Physicians

Jukka-Pekka “JP” Onnela

Department of Biostatistics
Harvard School of Public Health

April 27, 2013
• Wide variation in health care spending after adjusting for demographics
• Higher spending ≠ better outcomes
• Need to understand local medical practice style of physicians
• **Physicians’ direct professional connection to other physicians is through the patients they share**
• When physicians provide care to shared patients, they share clinical information through discussions and observations
• Leads to the formation of informal social networks
• These networks are organic in nature and may not conform to formal organizations
Use detailed Medicare records to quantify patient-physician interactions
- At the level of individual physician-physician ties
- At the level of groups of physicians that emerge over time

Ultimate goal to understand how networks impact quality and cost of care


1 Background

2 Network basics

3 From data to networks

4 Finding naturally occurring groups of physicians

5 Summary
Basic Definitions: Vertices, Edges, and Graphs

- **Graphs are mathematical models of network structures**
- We will use the terms network and graph interchangeably
- A graph is a way of specifying relationships among a collection of items
- Graphs consist of two kinds of elements:
  - Vertices (nodes)
  - Edges (ties, arcs)
A simple graph is an ordered pair \( G = (V, E) \).

Here \( V \) is the **vertex set** and \( E \) is the **edge set**.

The vertex set consists of vertices \( V = \{v_1, v_2, v_3, v_4\} \).

The edge set consists of pairs of vertices \( E = \{(v_1, v_2), (v_1, v_4), (v_2, v_4), (v_2, v_3)\} \).

Some vertex pairs are connected by an edge, some are not.

Two connected vertices are said to be (nearest) **neighbors**.
Connectivity of an undirected graph may be captured by an $N \times N$ symmetric adjacency matrix $A$ with entries

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \cdots & A_{NN} \end{pmatrix}$$

For a simple (unweighted, binary) graph

$$A_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are connected} \\ 0 & \text{otherwise} \end{cases}$$
The **degree** of a vertex in a graph is the number of edges connected to it
- We use $k_i$ to denote the degree of vertex $i$

If a graph is not connected, it breaks apart into **components**
- Each such disconnected piece consists of a group of nodes
- Each piece is connected when considered as a graph in isolation
- No two pieces overlap (no shared vertices)

![Diagram of a graph with disconnected components](image-url)
Basic Definitions: Communities

- Network **communities** correspond to tightly-knit regions of networks
- A working definition: “A group of nodes that are relatively densely connected to each other but relatively sparsely connected to other nodes in the network”
  - Social networks: core family, village, region, country, etc.
- Communities are thought to have a strong bearing on functional units in networks
Components vs. communities
1 Background

2 Network basics

3 From data to networks

4 Finding naturally occurring groups of physicians

5 Summary
Data for 100% of Medicare beneficiaries in 50 hospital referral regions (HRRs) from 2006, randomly sampled with probability proportional to their size

- Boston HRR was also included
- 4,586,044 Medicare beneficiaries seen by 68,288 physicians
- Included physician characteristics: age, sex, medical school, place of residency, zip code, and principal specialty (PCP, medical specialist, surgical specialist)
- Included: Medicare parts A and B
- Excluded: Patients enrolled in Medicare Advantage; specialties with no direct patient care (e.g., radiology, anesthesia); laboratory and other services not requiring a physician visit; physicians who saw fewer than 30 Medicare patients during 2006
The networks so far have been **unipartite or one-mode** networks.

The vertices in one-mode networks are all of the same type or kind.

In **bipartite or two-mode networks** there are two kinds of vertices, and the edges run only between vertices of different type.

Patient-physician encounter data can also be seen as bipartite graph.
Bipartite Graphs

- The equivalent of an adjacency matrix for a bipartite network is a rectangular matrix called an **incidence matrix**
- Consider a bipartite network of $N$ physicians and $M$ patients, where a physician and a patient are connected if the physician provides care to the patient
- Use the following notation to distinguish between the two types of vertices: $U = \{u_1, u_2, \ldots, u_M\}$ (patients) and $V = \{v_1, v_2, \ldots, v_N\}$ (physicians)
- For a simple unweighted bipartite network the incidence matrix $B$ is an $M \times N$ matrix having elements $B_{ij}$ such that

$$B_{ij} = \begin{cases} 
1 & \text{if vertex } u_i \in U \text{ and vertex } v_j \in V \text{ are connected} \\
0 & \text{otherwise}
\end{cases}$$
- It is often useful to work with direct connections between vertices of just one type.
- We can obtain a **one-mode projection** from the bipartite network.
- Consider vertex $u_1$ and edges $(u_1, v_1)$, $(u_1, v_2)$, and $(u_1, v_3)$.
- When projected to the $V$ vertex set, the vertices $\{v_1, v_2, v_3\}$ form a **clique**.
- Each vertex $u \in U$ that is connected to $n$ vertices in $V$ will result in an $n$-clique.

The projection, when taken over all $M$ vertices in $U$, will consist of the union of $M$ cliques, one for each vertex in $U$. 

![Diagram of Bipartite Graphs](image-url)
• Recall the definition of the incidence matrix $B_{ij}$:

$$B_{ij} = \begin{cases} 
1 & \text{if vertex } u_i \in U \text{ and vertex } v_j \in V \text{ are connected} \\
0 & \text{otherwise}
\end{cases}$$

• The product $B_{ki}B_{kj}$ will be 1 if and only if $v_i$ and $v_j$ are both connected to $u_k$ in the bipartite network.

• The number of vertices in $U$ to which both $v_i$ and $v_j$ are connected is

$$P_{ij} = P_{ji} = \sum_{k=1}^{M} B_{ki}B_{kj} = \sum_{k=1}^{M} B^T_{ik}B_{kj}$$

(2)

• The above projection can be also written in terms matrices as $P = B^T B$

• The element $P_{ij}$ encodes the “strength” of the connection between $v_i$ and $v_j$. 
Figure: Example of projecting a bipartite physician-patient network to a unipartite physician network. In this figure $B$ corresponds to $B^T$ in the text.
Descriptive Network Measures

- Start with 51 networks, one per HRR
- Keep only the largest connected component of each network
- The resulting networks tend to be very tightly connected; a small number of shared patients is unlikely to result in a meaningful information-sharing tie
- Filtered ties using **absolute thresholding** and **relative thresholding**
- Implemented relative thresholding maintaining 20% of ties for each physician
- No requirement for dyadic reciprocity
### Descriptive network measures

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Mean (SD) [Range]</th>
<th>Hospital Referral Region</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Boston, MA</td>
</tr>
<tr>
<td>Physicians</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No.</td>
<td>1339 (1621) [135-8197]</td>
<td>8197</td>
</tr>
<tr>
<td>Age, y</td>
<td>48.5 (1.3) [46.1-51.7]</td>
<td>48.3</td>
</tr>
<tr>
<td>Males, %</td>
<td>80.3 (4.9) [69.2-89.7]</td>
<td>70.1</td>
</tr>
<tr>
<td>Type of practice, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary care</td>
<td>41.9 (5.4) [27.6-53.3]</td>
<td>38.5</td>
</tr>
<tr>
<td>Medical specialist</td>
<td>30.0 (4.8) [20.7-47.1]</td>
<td>36.8</td>
</tr>
<tr>
<td>Surgical specialist</td>
<td>28.0 (3.3) [21.3-35.5]</td>
<td>24.7</td>
</tr>
<tr>
<td>Patients</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No.</td>
<td>279 (84) [126-447]</td>
<td>224</td>
</tr>
<tr>
<td>Whites, %</td>
<td>86.6 (11.5) [49.4-98.8]</td>
<td>89.1</td>
</tr>
<tr>
<td>Blacks, %</td>
<td>8.9 (10.1) [0.1-42.1]</td>
<td>6.2</td>
</tr>
<tr>
<td>Hispanics, %</td>
<td>1.8 (4.0) [0-24.2]</td>
<td>1.5</td>
</tr>
<tr>
<td>Females, %</td>
<td>59.9 (2.0) [54.4-64.2]</td>
<td>60.8</td>
</tr>
<tr>
<td>Medicaid, %</td>
<td>23.0 (10.0) [7.7-51.9]</td>
<td>27.7</td>
</tr>
<tr>
<td>Age, y</td>
<td>70.7 (1.6) [67.1-74.6]</td>
<td>70.8</td>
</tr>
<tr>
<td>Hierarchical clinical condition score</td>
<td>1.9 (0.3) [1.5-2.8]</td>
<td>2.2</td>
</tr>
<tr>
<td>Episode treatment group intensity score</td>
<td>1.03 (0.06) [0.9-1.2]</td>
<td>1.1</td>
</tr>
<tr>
<td>Networks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of ties</td>
<td>50 927 (75 963) [1568-392 582]</td>
<td>392 582</td>
</tr>
<tr>
<td>Adjusted degree&lt;sup&gt;a&lt;/sup&gt;</td>
<td>27.3 (10.4) [11.7-54.4]</td>
<td>51.4</td>
</tr>
<tr>
<td>No. of shared patients</td>
<td>852 (336) [297-1504]</td>
<td>835</td>
</tr>
<tr>
<td>Clustering</td>
<td>0.55 (0.06) [0.40-0.67]</td>
<td>0.48</td>
</tr>
<tr>
<td>Relative centrality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary care physician</td>
<td>0.38 (0.17) [0.19-1.06]</td>
<td>0.52</td>
</tr>
<tr>
<td>Medical specialist</td>
<td>3.47 (1.33) [0.48-7.40]</td>
<td>1.62</td>
</tr>
</tbody>
</table>

<sup>a</sup>Indicates the number of other physicians each physician was connected to per 100 Medicare beneficiaries across all hospital referral regions.
1 Background

2 Network basics

3 From data to networks

4 Finding naturally occurring groups of physicians

5 Summary
Many current policy changes intended to foster collaboration and improve care coordination among physicians.

Patient Protection and Affordable Care Act (2010) has a demonstration program for **Accountable Care Organizations (ACOs)**
- Groups of physicians providing concentrated care to their patients
- **Responsible for the total spending incurred by patients**
- Yet allows patients to go to any provider they choose

32 pilot ACOs began operation in early 2012

Organizations choosing to become ACOs assume substantial risk without tools to influence patient choices.
Current approaches to forming ACOs:
- Self-identification of physicians who will form an ACO
- Organize ACOs around single hospitals

Our approach:
- Form ACOs around groups of physicians who already have established patient-sharing patterns
Our approach

- Construct 51 networks, one for each HRR
- For each network, detect communities in the largest connected component
- Compare two ways of structuring ACOs:
  - A **community ACO** is identified with a network community
  - A **hospital ACO** is organized around a hospital and comprises the set of physicians for whom that hospital is their principal hospital
Clusters in networks

- A “network cluster” can be loosely defined as a “a group of nodes that are relatively densely connected to each other but relatively sparsely connected to other nodes in the network”
- Ability to discover clusters is a useful tool for revealing structure and organization within networks between microscopic and macroscopic levels

**Figure:** Homophily divides the school-based social network into similar groups. Here node color encodes race. Middle school and high school are separated into their own groups (top and bottom)

Source: EK Ref. 304.
Clusters in networks

- Need methods for finding clusters in large networks
- Two variants of the problem: **graph partitioning** and **community detection**
- Both refer to dividing the vertices of a network into groups
- In graph partitioning, the number and (approximate) size of the clusters is fixed
- In community detection, these are unspecified and need to be determined
• This problem comes up, for example, in parallel computing
• Consider simulating some complex dynamic process on a network
• We place variables on each vertex of the network
• To evolve the system, each vertex needs to have access to the variable associated with itself but also to the variables associated with the network neighbors
Clusters in networks: Graph partitioning

- In a two processor system, we would like to split the nodes evenly between the two processors.
- The main performance bottleneck is sending messages between the processors.
- We want to divide the graph into two parts (graph bisection) minimizing cut size, the number of edges between the groups.
- Exhaustive search is not an option since the number of ways of dividing a network such that $n_1 = n_2 = n/2$ is
  \[
  \binom{n}{n/2} = \frac{n!}{(n/2)!(n/2)!} \approx \frac{2^{n+1}}{\sqrt{n}}
  \]  

- There are several excellent heuristics for approaching this problem, such as the Kernighan-Lin algorithm.
• Communities are naturally occurring groups in networks regardless of their number or size
• In community detection, both the number of clusters and their sizes need to be determined
• Most algorithms result in non-verlapping partitions of vertices into groups
• Using minimum cut size is not helpful because that criterion is minimized by the trivial partition
• Need to reformulate the criterion or objective function
• A good division is one where there are fewer edges between communities than expected by chance
• Similar to graph partitioning, once the objective function has been specified, one needs heuristics to proceed
Community detection

- The most commonly used approach to community detection is modularity maximization, a global hard-partitioning method\(^1\) \(^2\)
- Let \(c_i\) denote the community (group) of vertex \(i\), where \(i = 1, 2, \ldots, n_c\); let \(m\) denote the number of edges
- Observed number of edges between vertices in the same community:
  \[
  \frac{1}{2} \sum_{i,j} A_{ij} \delta(c_i, c_j)
  \]
- Expected number of edges between vertices in the same community if edges are placed uniformly at random:
  \[
  \frac{1}{2} \sum_{i,j} \frac{k_i k_j}{2m} \delta(c_i, c_j)
  \]

\(^1\) Newman, PNAS 103, 8577 (2006)
\(^2\) Mucha, Richardson, Macon, Porter, Onnela, Science 328, 876 (2010)
Community detection

- Combining the observed and the expected counts:

\[
\frac{1}{2} \sum_{i,j} A_{ij} \delta(c_i, c_j) - \frac{1}{2} \sum_{i,j} \frac{k_i k_j}{2m} \delta(c_i, c_j) = \frac{1}{2} \sum_{i,j} \left( A_{ij} - \frac{k_i k_j}{2m} \right) \delta(c_i, c_j)
\]

\[
Q = \frac{1}{2m} \sum_{i,j} \left( A_{ij} - \frac{k_i k_j}{2m} \right) \delta(c_i, c_j) \in [-1, 1]
\]  

- Modularity maximization attempts to partition a network into communities in a way that maximizes the value of modularity
Table 1. Hospital Referral Region Characteristics (51 HRRs)

<table>
<thead>
<tr>
<th>Number of:</th>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beneficiaries</td>
<td>138,011</td>
<td>19,424</td>
<td>619,957</td>
</tr>
<tr>
<td>Physicians (Total)</td>
<td>1,339</td>
<td>135</td>
<td>8,197</td>
</tr>
<tr>
<td>Primary Care</td>
<td>545</td>
<td>68</td>
<td>3,158</td>
</tr>
<tr>
<td>Specialist</td>
<td>795</td>
<td>67</td>
<td>5,180</td>
</tr>
<tr>
<td>Hospital Networks*</td>
<td>18</td>
<td>2</td>
<td>45</td>
</tr>
<tr>
<td>Hospitals Networks+</td>
<td>8</td>
<td>1</td>
<td>36</td>
</tr>
<tr>
<td>Community Networks*</td>
<td>7</td>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td>Community Networks+</td>
<td>6</td>
<td>2</td>
<td>17</td>
</tr>
</tbody>
</table>

HRRs are as defined by the Dartmouth Atlas of Health Care
*Hospitals and Communities with 5 or more PCPs included
+Hospitals and Communities with at least 3,000 assigned patients
Figure: Networks in the Tallahassee, FL HRR based on hospital affiliation (A) and community affiliation (B).
### Table 2. Network Characteristics of Community and Hospital Networks

<table>
<thead>
<tr>
<th></th>
<th>Communities (n=273)</th>
<th>Hospitals (n=416)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Median (range)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Physicians</td>
<td>165 (21, 2470)</td>
<td>96*** (15, 1074)</td>
</tr>
<tr>
<td>% PCP (mean)</td>
<td>43 (19, 72)</td>
<td>44 (18, 79)</td>
</tr>
<tr>
<td>Number of Ties</td>
<td>4427 (104,253588)</td>
<td>2214*** (84, 71802)</td>
</tr>
<tr>
<td>Threshold</td>
<td>2 (1, 6)</td>
<td>2** (1, 8)</td>
</tr>
<tr>
<td>Adjusted Degree</td>
<td>27 (3, 202)</td>
<td>22*** (4, 176)</td>
</tr>
<tr>
<td>Physician Connections</td>
<td>626 (200, 1802)</td>
<td>572* (111, 1866)</td>
</tr>
<tr>
<td>% ties within the Network</td>
<td>50 (13, 99)</td>
<td>29*** (3, 94)</td>
</tr>
<tr>
<td>% patients shared within the Network</td>
<td>69 (17, 100)</td>
<td>49*** (7, 98)</td>
</tr>
<tr>
<td>Percent with at least 1:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orthopedist</td>
<td>97</td>
<td>97</td>
</tr>
<tr>
<td>Ophthalmologist</td>
<td>95</td>
<td>92</td>
</tr>
<tr>
<td>Cardiologist</td>
<td>96</td>
<td>87***</td>
</tr>
<tr>
<td>Neurologist</td>
<td>91</td>
<td>82**</td>
</tr>
<tr>
<td>Psychiatrist</td>
<td>84</td>
<td>76*</td>
</tr>
<tr>
<td>Dermatologist</td>
<td>85</td>
<td>75*</td>
</tr>
<tr>
<td>Gastroenterologist</td>
<td>86</td>
<td>82</td>
</tr>
</tbody>
</table>

Significance test compares the mean values.

*p<0.05
**p<0.01
***p<0.001
### Table 3. Percentage of Care in ACOs defined by Hospitals and Communities with at Least 5 PCPs and 3000+ Patients

<table>
<thead>
<tr>
<th></th>
<th>Mean % of care Provided in Networks Defined Based on:</th>
<th>Mean % of Care Provided in Network after adjusting for Network Size equal to the median sized hospital (n=96)</th>
<th>Mean % of Care Provided in Network after adjusting for Network Size equal to the median sized community (n=165)¹</th>
<th>Weighted Mean (range) % of care provided in Network across the 51 HRR Markets:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Admissions (one hospital per community)</td>
<td>Hospitals (n=416)</td>
<td>Communities (n=273)</td>
<td>Hospitals (n=416)</td>
<td>Communities (n=273)</td>
</tr>
<tr>
<td>Admissions (up to 2 hospitals)</td>
<td>66</td>
<td>73**</td>
<td>66</td>
<td>70##</td>
</tr>
<tr>
<td>Emergency Room Visits (one hospital per community)</td>
<td>37</td>
<td>41**</td>
<td>37</td>
<td>42###</td>
</tr>
<tr>
<td>Emergency Room Visits (up to 2 EDs)</td>
<td>--</td>
<td>75</td>
<td>--</td>
<td>79</td>
</tr>
<tr>
<td>Physician Visits</td>
<td>40</td>
<td>77***</td>
<td>39</td>
<td>74###</td>
</tr>
<tr>
<td>PCP Visits</td>
<td>72</td>
<td>91###</td>
<td>72</td>
<td>91###</td>
</tr>
<tr>
<td>Specialty Visits</td>
<td>8</td>
<td>61***</td>
<td>8</td>
<td>56###</td>
</tr>
</tbody>
</table>

*p<0.05  
**p<0.01  
***p<0.001  
###p<0.001

¹ Significance tests derived from the same regression model adjusted for number of physicians and the log number of physicians.
1 Background

2 Network basics

3 From data to networks

4 Finding naturally occurring groups of physicians

5 Summary
Summary

- Used community detection methods from network science to identify naturally occurring groups of physicians
- Communities encapsulate a large fraction of care, likely improving care coordination
- Our approach could be used to identify formal organizations that are most ready to become ACOs
Thanks!