Using Graphs to Characterize Nationwide Physician Referral Networks

Ding Tong
Yale University, ding.tong@yale.edu

Shu-Xia Li
Yale University, shu-xia.li@yale.edu

Isuru Ranasinghe
Yale University, isuru.ranasinghe@yale.edu

Sudhakar Nuti
Yale University, sudhakar.nuti@yale.edu

Hongyu Zhao
Yale University, hongyu.zhao@yale.edu

See next page for additional authors

Follow this and additional works at: http://elischolar.library.yale.edu/dayofdata

Part of the Biostatistics Commons, Health and Medical Administration Commons, and the Statistical Models Commons

http://elischolar.library.yale.edu/dayofdata/2014/Posters/16

This Event is brought to you for free and open access by EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Yale Day of Data by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.
Use Case Background
Treatment outcomes such as quality of care vary across hospitals. The drivers of such variations are not well understood. We hypothesize that the referral pattern between providers and the hospitals, can influence outcomes such as 30-day readmission. As a first step, we sought to use graph-based approaches to characterizing nationwide physician referral networks to better understand interactions among physicians. Our ultimate goal is to assess the association between hospital outcomes and their referral network characteristics.

Methods (continued)
2. Descriptive statistics of the physician referral networks at state level
For each state, we first calculated the characteristics of physicians including total number of physicians, and numbers of physicians in different specialties, physicians’ average graduation year, and gender composition. We then derived characteristics of networks, including the number of edges (referrals), average shared patients, average shared referral times, average betweenness (the number of shortest paths from all vertices to all others that pass through the node), primary care physician betweenness (average betweenness within primary care physicians) and the cluster coefficient. We visualized the primary care physician network with R software package igraph.

3. Statistical model for local network
After deriving the overall network characteristics, we used Exponential random graph models (ERGM) to further characterize local network property and sub-networks such as a physician referral sub-network of an individual hospital.

The ERGM model, we have:

\[ P(Y = y; \omega) = \exp(\gamma^T X - \omega^T \nu) \]

where \( Y \) is considered as a random matrix taking values in the set of all \( n \times n \) adjacency matrices. \( \gamma \) and \( \omega \) are vectors, \( X \) is a \( n \times n \) matrix (a nodal vector), \( \nu \) is a \( n \times n \) vector. \( Y \) is the normalized constant. \( P(Y; \omega) = \sum_y P(Y; \omega) \) where \( \omega \) is a function mapping \( 8^n \) into \( \omega \), is used to choose appropriate statistics.

Candidate covariates in the model include potentially clinically important factors and network property measures. Final model selection was based on statistical significance of the coefficients and akaike information criteria (AIC).

Traditional methods for assessing goodness of fit are not applicable to ERGM. To assess the goodness of fit, we first simulated network data 1000 times using the coefficients inferred from the selected model. To assess the overall goodness of fit, we compared the overall network properties between the observed network and the simulated networks. To assess the local goodness of fit, we generated the predicted network from the simulated networks by using a cutoff \( \text{Pr}(Y_{ij} = 1) \geq \frac{1}{n^2} \) for every dyad, where \( n \) is the number of nodes of edges of the observed network. We present the differences between the predicted and observed networks by heatmap organized according to covariates to detect local lack-of-fit.

Key Question
How to characterize local physician referral networks across the United States?

Novelty/Innovation
1. Applied graph-based evaluation approaches on nationwide physician referral network.
2. Developed a new method to evaluate goodness of fit for the statistical network models.

Results
1. Descriptive measures at state level: 1) graphs and network characteristics vary substantially across geographic areas, 2) graphs and characteristics show strong spatial correlations.

Results (continued)

Discussion
1. The descriptive network measures lack statistical rigor.
2. ERGM model can only use binary data instead of count data, and can not be easily adapted to deal with big sub-network.
3. The overall descriptive measures are at the state level but there is a need for studying smaller geographic regions.

Conclusions
Use of graph-based approaches has the ability to provide more insights in describing and evaluating the nationwide physician referral networks. In the future, we will refine the network descriptive measures, refit the ERGM model using both the Griffin-Hospital and nationwide network, and overcome the computational challenge of the ERGM and further adapting it for counting outcomes. We will further study the association between these network characteristics and health care outcomes.

References
[1] Krumholz HM et al. Goodness of fit of observed networks of physicians in a national dataset compared to a null model of network, and overcome the computational challenge of the ERGM and further adapting it for counting outcomes. We will further study the association between these network characteristics and health care outcomes.

Table 1. Data Sources and Variables

Table 2. Selected ERGM model output for a 27-Node sub-network of Griffin Hospital

Table 3. Goodness of fit: Observed vs. simulated networks respecting proportions of nodes vs. in- and out-degree, edge weight shared partner and Newman-Girvan modularity distance.

Appendix

Figure 1. a) Primary care physician density physicians per 1000 people.

Figure 2. Primary health care physician density (of physicians per 1000 people).

Figure 3. a) Goodness of fit: Observed network vs. predicted network by simulation. Left panel, white indicates existence of edges between observed and simulated, blue indicates the edges exist in observed but not simulated, red indicates otherwise; right panel shows the observed edges (value of 1 or 0) minus the empirical probability of the existence of such edges. The goodness of fit plots show no evidence for lack-of-fit.

Figure 4. a) Primary care physician referral network of Nevada.