

Improving Network Science in Health Services Research

Michael L. Barnett, MD, MS^{1,2}



¹Department of Health Policy and Management, Harvard T. H. Chan School of Public Health, Boston, MA, USA;

²Division of General Internal Medicine and Primary Care, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA.

J Gen Intern Med

DOI: 10.1007/s11606-019-05264-4

© Society of General Internal Medicine 2019

In the past 20 years, the discipline of network science has grown at a breathtaking pace. Its influence has touched nearly every domain of scientific inquiry, including genetics, neuroscience, computer science, economics, and epidemiology, to name just a few.¹ The appeal of network science comes from a universal framework of tools to define nearly any complex system as a combination of “nodes” (a single point) and “edges” (a connection between two points). Studying networks has never been more accessible due to the increasing availability of free courses and textbooks, faster and cheaper computers, sophisticated visualization techniques, and powerful open source software. In 2019, it could be argued that every social scientist should have at least some familiarity with network methods.

It is no surprise that health services research has begun to embrace network science.² Network methods offer a completely distinct set of tools from standard epidemiological and statistical techniques to understand the complex and interconnected health care delivery network in the USA. Health services research has deep roots in the early days of network analysis. Some of the earliest pioneering studies in social network analysis examined the diffusion of new pharmaceuticals among physicians in the USA in the 1950s.³ Health services researchers are particularly fortunate to have a wealth of data that can be examined with a network mindset. For example, a state hospital discharge database is not just a record of individual inpatient admissions. It's also a complex network of hospitals connected by patient transfers. In the same way, an insurance claims database is not just a list of health care services by clinicians, it's a network of clinicians and facilities, all connected by the flow of patients.

We have already learned a great deal about health care delivery through the network science approach. One fundamental observation is that networks of health care providers are highly variable, with a wide range of “centrality” (connectedness to the entire network) of primary care and specialist physicians within regions, and a strong association between primary care centrality and lower costs.^{4, 5} There is growing evidence that patients treated by broader, more disconnected

teams or networks of physicians have higher costs, higher rates of admissions and emergency department visits, and lower quality care.^{4, 6, 7} We can also observe how doctors influence each other's practice along networks in areas such as prescribing and imaging, with important implications for understanding the spread of new technologies.^{3, 8} Another exciting application is that networks can be used to study the dynamics of nosocomial infection between health care facilities.⁹ This set of contributions is impressive, but much of the research above has relied on insurance claims data from one data source, often Medicare.

This is the context necessary to appreciate the network study by Dr. Trogdon and colleagues¹⁰ in this month's issue of *JGIM*. Their study aims to address a fundamental challenge in studying many networks, which is missing data. For networks in health services research, the biggest challenge is that few data sources capture the entirety of patient care for a single physician, and ever fewer at the scale of an entire state or the whole country. For example, Medicare administrative claims data are incredibly detailed, but any physician patient-sharing networks built using these data completely miss any connections between physicians who share patients not covered by Medicare. This problem gets even more complicated when using randomly sampled datasets, such as 20% sample files, that are frequently the most accessible for researchers. So, given a dataset from just one insurer, how much information are we losing?

The answer is not clear yet, but Trogdon and colleagues make an important dent in the problem. They use a dataset of colon cancer registry patients in North Carolina linked to claims from all of the major insurance types: private, Medicare and Medicaid, from 2003 to 2013. These data, while only focused on a narrow subset of patients, can give us insight into how much we miss by studying networks with just data from just one payer. They use a standard approach to constructing networks using data from Medicare and private insurance claims together (Medicaid data could not be included due to technical limitations). They then looked at connections between surgeons, medical oncologists, and radiation oncologists who share 2 or more patients over this 10-year period. They calculate a wide range of network statistics, including more sophisticated “community detection” methods, which can find “clumps” of physicians who are more connected with each other than with physicians outside the group.

Their most fundamental observation is that only 31% of all of the connections between two physicians appeared in both networks built using Medicare claims data or private insurance claims. This is certainly lower than I would have hoped, but also not entirely surprising given the significant differences between these populations. The major culprit for missing links was the private insurance network, which missed over half of the connections seen in the complete network with both payers' data. Not surprisingly, missing half of these connections meant that there was a very poor correlation between the Medicare and private insurance-based networks for measures that rely heavily on the detailed structure of the network such as centrality. Also, the communities of physicians detected in the networks made from each payer's claims had important differences, though there is no standard approach to quantify whether this difference is large or small.

On the other hand, several important measures were surprisingly similar. One of the most fundamental network measures is "degree," or the number of connections each physician has. Even though the authors described degree as "depending heavily" on the payer, the average degree in networks made using private claims or a similar sized sample of Medicare claims was 26.0 and 27.2, respectively. To me, this is more notable for how close the data sources are despite only sharing 31% of connections. Similarly, important measures of local connectivity between small groups of physicians called "clustering coefficients" were surprisingly close using either dataset. Though these similarities are reassuring, Figure 2 in Trogdon et. al shows that the similar average values mask a significant amount of discordance in measures for many individual physicians.

There are important limitations that qualify interpretation of this study. The most significant is that their networks are limited to a modest sample of patients and only 3 specialties. On top of that, these patients are in one state and the authors were unable to use Medicaid claims in this analysis due to data limitations. All of this raises the question of whether this sample of patients and specialists produces different results than we would see in a much broader dataset or a different clinical context. In a way, the study raises the meta-question of how we should interpret their own results.

In the end, this analysis makes me optimistic for the continued growth of network methods in our field. Trogdon and

colleagues provide an insightful analysis that pushes us to closely consider our use of these new tools. The cycle of learning how to address the limitations of prior research and strengthening our methods is how we can advance the quality and impact of health services research.

Corresponding Author: Michael L. Barnett, MD, MS; Department of Health Policy and Management Harvard T. H. Chan School of Public Health, Boston, MA, USA; (e-mail: mbarnett@hsph.harvard.edu).

Compliance with Ethical Standards:

Conflict of Interest: The author declares that he does not have a conflict of interest.

REFERENCES

1. **Barabási A-L, Posfai M.** Network Science. Cambridge University Press; 2016.
2. **DuGoff EH, Fernandes-Taylor S, Weissman GE, Huntley JH, Pollack CE.** A scoping review of patient-sharing network studies using administrative data. *Transl Behav Med.* 2018;8(4):598–625. <https://doi.org/10.1093/tbm/ibx015>.
3. **Coleman J, Katz E, Menzel H.** The diffusion of innovations among physicians. *Sociometry.* 1957;20(4):253–270.
4. **Barnett ML, Christakis NA, O'Malley J, Onnela J-P, Keating NL, Landon BE.** Physician Patient-sharing Networks and the Cost and Intensity of Care in US Hospitals. *Med Care.* 2012;50(2):152.
5. **Landon B, Keating N, Barnett M, et al.** Variation in patient-sharing networks of physicians across the United States. *JAMA J Am Med Assoc.* 2012;308(3):265.
6. **Landon BE, Keating NL, Onnela J-P, Zaslavsky AM, Christakis NA, O'Malley AJ.** Patient-Sharing Networks of Physicians and Health Care Utilization and Spending Among Medicare Beneficiaries. *JAMA Intern Med.* 2018;178(1):66–73. <https://doi.org/10.1001/jamainternmed.2017.5034>.
7. **Agha L, Ericson KM, Geissler KH, Rebitzer JB.** Team Formation and Performance: Evidence from Healthcare Referral Networks. National Bureau of Economic Research; 2018. <https://doi.org/10.3386/w24338>.
8. **Pollack CE, Soulos PR, Herrin J, et al.** The Impact of Social Contagion on Physician Adoption of Advanced Imaging Tests in Breast Cancer. *J Natl Cancer Inst.* 2017;109(8). <https://doi.org/10.1093/jnci/djw330>
9. **Fernández-Gracia J, Onnela J-P, Barnett ML, Eguíluz VM, Christakis NA.** Spread of pathogens in the patient transfer network of US hospitals. In: International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation. Springer, Cham; 2017:271–280.
10. **Trogdon JG, Weir WH, Shai S, et al.** *J Gen Intern Med.* 2019. <https://doi.org/10.1007/s11606-019-04978-9>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.