Variation in Patient-Sharing Networks of Physicians Across the United States

Bruce E. Landon, M.D., M.B.A[1,2], **Nancy L. Keating, M.D., M.P.H.**[1,3], **Michael L. Barnett, M.D.**[1,3], **Jukka-Pekka Onnela, Ph. D.**[4], **Sudeshna Paul, Ph.D.**[1], **A. James O'Malley, Ph.D.**[1], **Thomas Keegan, Ph.D.**[1], and **Nicholas A. Christakis, M.D., Ph.D.**[1,2,5]

[1]Department of Health Care Policy, Harvard Medical School, Boston, MA

[2]Division of Primary Care and General Internal Medicine, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA

[3]Division of General Internal Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, MA

[4]Department of Biostatistics, Harvard School of Public Health, Boston, MA

[5]Department of Sociology, Harvard University, Cambridge, MA

Abstract

Context—Physicians are embedded in informal networks that result in their sharing patients, information, and behaviors.

Objective—We use novel methods to identify professional networks among physicians, to examine how such networks vary across geographic regions, and to determine factors associated with physician connections.

Design, Setting, and Participants—Using methods adopted from social network analysis, we used Medicare administrative data from 2006 to study 4,586,044 Medicare beneficiaries seen by 68,288 active physicians practicing in 51 hospital referral regions (HRRs). Distinct networks depicting connections between physicians (defined based on shared patients) were constructed for each of the 51 HRRs.

Main Outcomes Measures—Variation in network characteristics across HRRs and factors associated with physicians being connected.

Results—The number of physicians per HRR ranged from 135 in Minot, ND to 8,197 in Boston, MA. There was substantial variation in network characteristics across HRRs. For example, the median adjusted degree (number of other physicians each physician was connected to, per 100 Medicare beneficiaries) across all HRRs was 27.3, and ranged from 11.7 to 54.4; also, primary care physician (PCP) relative centrality (how central PCPs were in the network relative to other physicians) ranged from 0.19 to 1.06. Physicians with ties to each other were far more likely to be based at the same hospital (96.0% of connected physician pairs versus 69.2% of unconnected pairs, p<.001, adjusted rate ratio=.12, 95% confidence interval .12,.12)), and were in closer geographic proximity (mean office distance of 13.2 miles for those with connections versus 24.2 for those without, p<.001). Connected physicians also had more similar patient panels, in terms of

Address correspondence to Bruce E. Landon, Department of Health Care Policy, Harvard Medical School, 180 Longwood Avenue, Boston, MA 02215. Ph: 617-432-3456, Fax: 617-432-0173; landon@hcp.med.harvard.edu.

the race or illness burden, than unconnected physicians. For instance, connected physician pairs had an average difference of 8.8 points in the percentage of black patients in their two patient panels compared with a difference of 14.0 percentage points for unconnected physician pairs (p<. 001).

Conclusions—Network characteristics vary across geographic areas. Physicians tend to share patients with other physicians with similar physician-level and patient-panel characteristics.

Keywords

Physician behavior; physician networks; Geographic variation; social networks

Introduction

In 1973, Wennberg and Gittlesohn first described the extent to which local practice patterns varied across towns in Vermont.[1] Decades of subsequent research demonstrating both small- and large-area variations in care suggest that local norms play an important role in determining practice patterns and that, in aggregate, such norms and customs might account for a large proportion of the variability that exists in health care.[2–4] Whatever the exact cause, small-area variation in patterns of care suggests that physicians may come to conform to the behavior of other, nearby physicians.

In part, this might happen by physicians actively sharing clinical information among themselves through informal discussions and observations (e.g., of patient records) that occur in the process of providing care to shared patients.[5] These informal information-sharing networks of physicians differ from formal organizational structures (such as a physician group associated with a health plan, hospital, or independent practice association) in that they do not necessarily conform to the boundaries established by formal organizations, although such formal organizational affiliations clearly influence the interactions physician have. Informal information-sharing networks among physicians may be seen as organic or natural rather than as artificial or deliberate.

The potential influence of informal networks of physicians on physician decision-making has been neglected despite the potential importance of these networks in day-to-day practice. In addition, understanding more about physicians' predilections to form relationships with colleagues could be important for identifying levers to influence how physicians exchange information with one another. Here, we use novel, validated methods based on patient sharing to define professional networks among physicians, and we examine how such networks vary across geographic areas. We also identify physician and patient-population factors that are associated with patient-sharing relationships.

Methods

Overview

Sharing of patients based on administrative data can identify information-sharing ties among pairs of physicians.[6] We use physician encounter data from the Medicare program to define networks of physicians based on shared patients.[7] A social network is defined as a set of actors and the relationships or connections that link these actors together. Social network analysis can be used to study the structure of a social system and to understand how this structure influences the behavior of constituent actors. As we define networks in the present application, nodes represent the individual physicians in the network and ties (or edges) represent shared patients between nodes. We use the presence of shared patients to infer information-sharing relationships between two physicians. Ties vary in their "weight"

according to the number of shared patients, with more shared patients implying stronger connections between physicians.[6]

Identifying the Sharing of Patients

Shared patients were identified using Medicare claims from 2006. To maximize data on shared patients among physicians practicing in local areas, we obtained data for 100% of Medicare beneficiaries (including those under age 65) living in 50 market areas (defined as hospital referral regions, or HRRs) randomly sampled with probability proportional to their size; this was the maximum amount of data CMS would release. HRRs represent regional markets for tertiary care defined based on cardiovascular and neurosurgical procedures.[8] In addition, the Boston HRR was included to aid in the development and testing of our methods since it is familiar to us. We included in our analyses patients enrolled in both Medicare Parts A and B. We excluded patients enrolled in Medicare Advantage plans, for whom encounter data are not available.

We defined encounters with physicians based on paid claims in the carrier file. We excluded claims for non-direct patient care specialties or specialties where individual physicians are not selected (e.g., anesthesia, radiology). We identified all evaluation and management services, and also included procedures with a relative value unit (RVU) value of at least 2.0 in order to capture surgical procedures that often are reimbursed via bundled fees that include pre- and post-procedure assessments. We excluded claims for laboratory and other services not requiring a physician visit; we also excluded claims generated from physicians who saw fewer than 30 Medicare patients during 2006 or who practice outside of the included HRRs. Although the latter exclusions risk dropping physicians who work on the geographic boundary of a HRR, information on these physicians would have been incomplete.

Physician and patient population characteristics

We defined physician characteristics including age, sex, medical school, and place of residency using data from the AMA Physician Masterfile.[9] We used billing zip code and specialty designation from the Medicare claims (defined based on the plurality of submitted claims) to assign a principal specialty and practice location. We excluded physicians (<1%) for whom we could not identify a dominant specialty or practice location. We classified physicians as primary care physicians (PCPs, defined as general internists, family practitioners, or general practitioners), medical specialists, or surgical specialists. For each physician, we calculated the following practice-level variables for their Medicare patients: mean age, percent female, race/ethnicity composition (% white, black, and Hispanic), and mean health status measured using the Centers for Medicare and Medicaid Services Hierarchical Condition Categories (HCC) risk adjustment model.[10] We also characterized each physician's practice style based on the intensity of care delivered to their patients, measured using Episode Treatment Group software.[11] We defined intensity of care based on the mean resource use for episodes of care delivered by that physician compared with similar episodes delivered by all other physicians in our data (see the Technical Appendix for more details). A physician with an intensity index of 1.0 delivers care for all the conditions she treats that is equal to the intensity (as measured by total service use) of the average physician of the same specialty in our data treating those conditions. A score of 1.2 would indicate that she is 20% more costly.

Constructing Physician Networks

The structural backbone from which we discern physician networks is a patient-physician "bipartite" or two-mode network. The term "bipartite" means that nodes in the network can be partitioned into two sets, physicians and patients, and that all relationships link nodes

from one set to the other.[12] We form a unipartitite (physician-physician) network[13,14] by connecting each pair of physicians who share patients with one another, and we assign a weight to such ties based on the number of patients shared (Figure 1 provides additional details). A key decision involved determining the minimum number of shared patients that could optimally be used to define connections representing important relationships. We previously found that physicians in a single academic health care system who shared eight or more patients using these same Medicare data had an 80% probability of having a validated information-sharing relationship.[6] This threshold might differ depending upon specialty and the clinical activity of each individual physician. We explored both *absolute thresholding* (using the same threshold for each physician and specialty) and *relative thresholding* (creating a customized threshold for each physician). We found that using a relative threshold that maintained 20% of the strongest ties for each physician appeared to best maintain intrinsic network characteristics while also eliminating noise that might result from spurious connections. Although this method likely eliminates some ties that represent true relationships, it maintains the strongest ties for each physician and therefore maintains the relationships likely to be most influential to that physician. In sensitivity analyses using the top 10% and the top 30% of ties, our main results were similar (please see the Technical Appendix for additional details).

Network Descriptive Measures

We described the networks (after applying our thresholding procedure) focusing on a set of network measures applicable across all types of physicians: adjusted physician degree, number of patients shared by the physician, relative betweenness centrality, and physician-level clustering coefficient (see Figure 2).

Degree is defined as the number of doctors connected to a given physician through patient sharing. Because the number of connections is influenced by patient volume, we adjusted degree by dividing each physician's degree by the total number of Medicare patients the physician shared with other doctors (*adjusted degree)*. Thus, physicians with a higher degree are connected to, and share patients with, more physicians. The *number of shared patients* is the total number of shared patients across *all* ties for an individual physician and reflects the number of patients that physician cares for, as well as his tendency to share care with other physicians. The *betweenness centrality* of a physician represents how central a physician is within his network of colleagues.[14–16] To calculate relative *betweenness centrality* for PCPs or medical specialists in each HRR-level network, we calculate mean PCP or specialist centrality for that network and divide by the mean centrality of all other physicians in the network. Central physicians in a network are likely to have more influence. The *clustering* coefficient of a physician in the network refers to the proportion of a physician's colleagues who also share patients with one another. A doctor could share patients with 10 other doctors, none of whom share patients with each other, or a doctor could share patients with 10 other doctors all of whom are interconnected. A network with a high clustering coefficient is more densely connected.

For descriptive purposes, we assigned each physician to a hospital. Those whose practice address was located in a sampled HRR were assigned to a principal hospital based on where they filed the plurality of inpatient claims or, if they did not do inpatient work, to the hospital where the plurality of patients they saw received inpatient care.[17]

Statistical Analyses

We first characterized the networks in each of the 51 HRRs. Selected networks are visualized using the Fruchterman-Rheingold algorithm.[18] We assessed unadjusted differences in network measures across regions using one-way analysis of variance.

To examine factors associated with the existence of ties between physicians, we first compared the characteristics of pairs of connected physicians within each of the regions to the characteristics of all other potential pairs where there was no connection. These analyses included *all* physicians and ties (not just the 20% of strongest ties used for the descriptive network analyses). For each physician pair, whether connected or not, we defined difference measures for each of the main independent variables of interest. For instance, distance was defined as the number of miles between the zip code centroid for each pair's office addresses (and was log-transformed to limit the impact of outliers). We excluded shared patients from the calculation of patient-panel attributes for each physician pair, so our results are not inflated by the fact that shared patients have identical characteristics.

Bivariate differences were evaluated using two-sided t-tests or chi-square tests, as appropriate, and were considered significant if the p-value was less than .05. Because our analyses are hypothesis-generating, we did not adjust p-values for multiple comparisons. We then estimated univariable and multivariable models to identify characteristics associated with having a tie and increasing tie strength between two physicians within the network (i.e., the extent to which characteristics of two doctors connected to each other are similar, also known as homophily). The dependent variable was the number of shared patients between any pair of physicians and the predictors were the difference measures as shown in Table 2. Because the prevalence of potential ties with zero patients was large, we found that a negative binomial distribution fit the data the best. To make the results easier to interpret, regression coefficients were converted to standard rate ratios because the outcome is not binary. For the differences in patient characteristics measured as percentages, we present rate ratios representing the increase in number of ties for each 10% difference in a patient population characteristic across the two physicians. All analyses were performed with SAS, version 9.2 (Cary, NC).[19] This study was approved by the Institutional Review Board of Harvard Medical School, which also approved a waiver of consent for participants in the study.

Results

We studied 4,586,044 Medicare beneficiaries from 51 HRRs who were seen by 68,288 physicians practicing in those HRRs. The randomly sampled HRRs are distributed across all regions of the country and include urban and rural locations (Appendix Figure A1). The characteristics of all included physicians and patients are presented in Table 1. The mean physician age was 48.8 years and about 80% were male. Among the Medicare patients, the mean age was 71 and 40% were male. The distribution of the number of shared patients between linked physicians for the entire dataset is depicted in Appendix Figure A2.

After applying the relative thresholding rule (keeping only ties with strength in the top 20th percentile for each physician), the mean number of patients shared per 100 Medicare beneficiaries across the entire sample was 27.3. Network attributes are depicted graphically in Appendix Figure A3, which shows scatter plots of the network topological characteristics of interest versus the network size for the included HRRs. The network measures fall into two distinct categories: those with a strong dependence on network size (adjusted degree, clustering, number of shared patients) and those less associated with network size (relative PCP and medical specialist centrality). Network characteristics across the geographic regions also are shown in Table 1. Substantial variation was observed across HRRs. For example, the number of included physicians ranged from 135 in Minot, ND (with 1,568 ties) to 8,197 in Boston (392,582 ties). Physician adjusted degree is much higher in Boston (the average physician was connected to 51.4 other physicians per 100 Medicare patients cared for in Boston versus 11.7 in Minot), whereas clustering is greater in Minot (.62 in Minot versus .48 in Boston -- the clustering coefficient ranges from 0–1 and quantifies the

proportion of physicians who, in addition to being connected to a given physician, are also connected to one another). As noted above, these network characteristics also were strongly associated with network size. Other variation cannot be explained by the general relationship to network size, however, such as the greater relative betweenness centrality of specialists in Minot vs. Boston (Specialists are over 5 times more central than PCPs in Minot whereas in Boston they are only 1.6 times as central), meaning that certain structural aspects of physician networks are not simply functions of network size.

Graphical depictions of networks for two HRR's are presented in Figure 3. For descriptive purposes, hospital affiliation and specialty are presented in two separate graphs. In St. Paul, MN, there are many ties between physicians in different hospitals, with primary care physicians centering their patient sharing around a pool of medical and surgical specialists in multiple hospitals. Alternatively, in Albuquerque, NM, network connections are mostly confined within hospitals, and connections are generally confined to their hospital.

Graphical depictions of networks for two HRR's are presented in Figure 3. Networks are pictorially represented using "spring embedder" methods, which position objects with stronger connections (i.e., physicians with more shared patients) in closer physical proximity within the network. In St. Paul, MN, there are many ties between physicians in different hospitals, with primary care physicians centering their patient sharing around a pool of medical and surgical specialists in multiple hospitals. Thus, although physicians are clustered according to their principal hospital affiliation, the close proximity of the clusters is indicative of multiple ties across hospitals. Alternatively, in Albuquerque, NM, network connections are mostly confined within hospitals, and connections are generally confined to their hospital. Consequently, the hospital clusters in Albuquerque are more distinct and separated in space.

Factors Associated with Network Ties

Among all physicians and ties (rather than just the 20% of strongest ties), across the 51 HRRs, male physicians were more likely to have ties with other male physicians (65.1% of connected physician pairs were male-male versus 54.6% of unconnected physician pairs, p<.001), but female physicians were less likely to have ties with other female physicians (3.8% of connected physician pairs versus 6.4% of unconnected physician pairs, p<.001) (Table 2). Physicians with ties were also closer in age (mean difference of 11.5 years for those with ties versus 12.5 for those without, p<.001). Patterns varied by physician specialty as well. Although most (69.2%) connected physician pairs were from different hospitals, virtually all unconnected physician pairs (96.0%) were from different hospitals (p<.001). Connected physician pairs were also more likely to be in close geographic proximity. The mean distance for connected pairs was 13.2 miles versus 24.2 miles for unconnected pairs (p<.001). Connected physicians also had more similar practice intensity as measured by ETGs (a difference of .29 for linked physicians versus a difference of .31 for unlinked physicians, p<.001).

Characteristics of physicians' patient populations were also associated with the presence of ties between physicians. Across all physician racial and ethnic groups, connected physicians had more similar racial composition of their patient panels (net of any shared patients) than unconnected physicians. For instance, connected physician pairs had an average difference of 8.8 points in the percentage of black patients in their two patient panels compared with a difference of 14.0 percentage points for unconnected physician pairs (p<.001). Similarly, differences in mean patient age and percent Medicaid patients were also smaller for connected physicians than unconnected physicians. Medical comorbidities (measured by the HCC score) were also more similar, suggesting that connected physicians had more similar

patients in terms of clinical complexity than unconnected physicians. All of these results were confirmed in multivariable regression models as shown in the right column of Table 2.

Physicians thus tend to cluster together along attributes that characterize their own backgrounds and the clinical circumstances of their patients. Of note, we observed similar patterns when repeating the analyses using logistic regression after applying the thresholding criteria.

Discussion

Our results demonstrate substantial variation in physician network characteristics across geographic areas in terms of both topology and dyadic ties, even for networks of generally similar size. It has long been known that physician behavior varies across geographic areas, yet our understanding of the factors that contribute to these geographic differences is incomplete.[1] Our findings suggest that variation according to network attributes might help explain health care variation across geographic areas, particularly given what is known about how networks function. However, additional studies are needed to ascertain the extent to which the structural variation in physician interactions–on both macro and micro scales– can help explain variation in medical practice across geographic areas.

Our results show that physicians tend to share patients with colleagues who have similar personal traits, practice styles, and patient panels, although the influence of some of these traits is small in magnitude. Working at the same primary hospital and having similar sociodemographic characteristics among patients in their patient panel increase the likelihood of a connection between doctors. This extends prior research by incorporating measures of the strength of the tie and by examining predictors of ties.[20] Our work also extends prior research focused on restricted types of care (e.g. intensive care) or doctors (e.g., urologists).[21,22]

The network interactions among physicians that we discerned differ from the formal networks sometimes established by health plans or health systems because they reflect actual patient flows across physicians. Formal networks are clearly important, as evinced by the unsurprising finding that physicians associated with the same hospital are far more likely than other physicians to be connected. Yet this is not always the case. For instance, although hospital affiliation appears to be a strong predictor of ties in the Albuquerque market, this is not the case in St. Paul.

Our ability to discern these "organic," natural networks is relevant to the current push towards the creation of accountable care organizations (ACOs). Here, we define and identify natural groupings of physicians who are already sharing patients to provide care. Such doctors have a history of working with each other, and likely have evolved natural communication channels. Insurers and policy makers might therefore find it more efficient to identify candidate ACOs in this fashion.

When asked, physicians almost uniformly report that they choose other physicians for advice or referrals because of their skill and clinical expertise.[23–25] Clearly, however, physician associations are more complex and are related to other factors as well. Physicians demonstrate homophily in their professional networks just as in virtually every other social circumstance studied.[26–28] Because our data are cross-sectional, we cannot tell whether physicians preferentially choose to refer to physicians who treat similar patients or whether these physicians share similar patient populations for other reasons. It is notable, however, that these findings hold even when accounting for hospital affiliation. The extent of homophily we observe has additional implications: it might reduce the diffusion of valuable or novel information, and it could also increase the consistency of clinical practice. That is,

to the extent that physician are connected to doctors they already resemble, they are less likely to be exposed to novel information.

These analyses are subject to several limitations. First, we used Medicare data to identify shared patients among physicians. Patterns of patient-sharing may differ for younger patients or patients in integrated delivery systems. The growing availability of all-payer databases at the state level should facilitate more complete ascertainment of physician networks, although the Medicare data continue to have the advantage that they are available across the entire country. Second, our data were limited to a single year. Future analyses should replicate our findings using multiple years of data and examine the stability of networks over time. Third, our analyses of network characteristics included physician ties only if they were in the top 20% of ties for each individual physician. Some ties that we eliminated were likely to be true information-sharing relationships and conversely some that we retained may be ad hoc or happenstance (though our sensitivity analyses confirmed the robustness of the findings). Moreover, our approach fails to capture physician interactions with other physicians across the country through professional societies and likely underestimates information sharing among physicians within specialty. Fourth, our model of characteristics associated with ties assumes conditional independence of dyads; currently available statistical methods precluded accounting for potential network-generated dependencies in datasets of our size. Fifth, the rapid adoption of electronic medical records since 2006[29] could lead to different relationship patterns. Finally, although we demonstrate variation in networks across geographic areas, additional research is needed to establish whether network characteristics are associated with variations in care.

In conclusion, we used novel methods to define social networks among physicians in geographic areas based on shared patients, examined how such networks vary across different geographic regions, and identified physician and patient-population factors that are associated with physician patient-sharing relationships. We believe that our approach might have useful applications for policymakers seeking to influence physician behavior (whether related to ACO formation of innovation adoptions) as it is likely that physicians are strongly influenced by their network of relationships with other physicians.

Acknowledgments

References

1. Wennberg J, Gittelsohn. Small area variations in health care delivery. Science. 1973; 182(117): 1102–1108. [PubMed: 4750608]

2. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The Implications of Regional Variations in Medicare Spending. Part 1: The Content, Quality, and Accessibility of Care. Ann Intern Med. 2003; 138(4):273–288. [PubMed: 12585825]

3. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The Implications of Regional Variations in Medicare Spending. Part 2: Health Outcomes and Satisfaction with Care. Ann Intern Med. 2003; 138(4):288–299. [PubMed: 12585826]

4. Soumerai SB, McLaughlin TJ, Gurwitz JH, et al. Effect of local medical opinion leaders on quality of care for acute myocardial infarction: a randomized controlled trial. Jama. 1998; 279(17):1358–1363. [PubMed: 9582043]

5. Coleman, JS.; Katz, E.; Menzel, H. Medical Innovation: A Diffusion Study. Indianapolis, IN: Bobbs-Merrill; 1966.

6. Barnett ML, Landon BE, O'Malley AJ, Keating NL, Christakis NA. Mapping physician networks with self-reported and administrative data. Health Serv Res. 2011; 46(5):1592–1609. [PubMed: 21521213]

7. 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):152–160. [PubMed: 22249922]

8. Dartmouth Medical School Center for the Evaluative Clinical Sciences. Dartmouth Atlas of Health Care. American Hospital Publishing, Inc; 1998.

9. Baldwin L-M, Adamache W, Klabunde CN, Kenward K, Dahlman C, Warren JL. Linking physician characteristics and Medicare claims data: issues in data availability, quality, and measurement. Med Care. 2002; 40(8 supp):IV82–IV95.

10. Pope, GC.; Kautter, J.; Ellis, RP., et al. [Accessed December 17, 2010.] Health Care Financing Review: Risk Adjustment of Medicare Capitation Payments Using the CMS-HCC Mode. 2004. http://cms.gov/HealthCareFinancingReview/downloads/04Summerpg119.pdf

11. Ingenix. [Accessed 7/13/11, 2011.] What are ETGs?. 2008. http://www.ingenix.com/content/File/What_are_ETG.pdf

12. Everett, MG.; Borgatti, SP. Extending Centrality. In: Carrington, PJ.; Scott, J.; Wasserman, S., editors. Models and Methods in Social Network Analysis. New York: Cambridge University Press; 2005.

13. Breiger RL. The Duality of Persons and Groups. Social Forces. 1974; 53(2):181–190.

14. Wasserman, S.; Faust, K. Network Analysis: Methods and Applications. New York: Cambridge University Press; 1994.

15. Keating NL, Ayanian JZ, Cleary PD, Marsden PV. Factors affecting influential discussions among physicians: a social network analysis of primary care practice. J Gen Intern Med. 2007; 22(6):794–798. [PubMed: 17404798]

16. Newman, M. Networks: An Introduction. 1. Oxford: Oxford University Press; 2010.

17. Bynum JP, Bernal-Delgado E, Gottlieb D, Fisher E. Assigning ambulatory patients and their physicians to hospitals: a method for obtaining population-based provider performance measurements. Health services research. Feb; 2007 42(1 Pt 1):45–62. [PubMed: 17355581]

18. Fruchterman TMJ, Reingold EM. Graph drawing by force-directed placement. Software-Practice and Experience. 1991; 21(11):1129–1164.

19. SAS Institute I, SAS. [Accessed 7/12/11, 2011.] 2011. http://www.sas.com/

20. Pham HH, O'Malley AS, Bach PB, Saiontz-Martinez C, Schrag D. Primary care physicians' links to other physicians through Medicare patients: the scope of care coordination. Ann Int Med. 2009; 150:236–242. [PubMed: 19221375]

21. Pollack CE, Weissman G, Bekelman J, Liao K, Armstrong K. Physician social networks and variation in prostate cancer treatment in three cities. Health Serv Res. Feb; 2012 47(1 Pt 2):380–403. [PubMed: 22092259]

22. Iwashyna TJ, Christie JD, Kahn JM, Asch DA. Uncharted paths: hospital networks in critical care. Chest. Mar; 2009 135(3):827–833. [PubMed: 19265091]

23. Kinchen KS, Coper LA, Levine D, Wang NY, Powe NR. Referral of patients to specialists: factors affecting choice of specialist by primary care physicians. Ann Fam Med. 2004; 2(3):245–252. [PubMed: 15209202]

24. Barnett M, Keating N, Christakis N, O'Malley A, Landon B. Referral patterns and reasons for referral among primary care and specialist physicians. J Gen Intern Med. 2011 Epub Sept. 16.

25. Keating NL, Zaslavsky AM, Ayanian JZ. Physicians' experiences and beliefs regarding informal consultation. Jama. 1998; 280:900–904. [PubMed: 9739974]

26. McPherson M, Smith-Lovin L, Cook JM. Birds of a feather: homophily in social networks. Annual Review of Sociology. 2001; 27(1):415–444.

27. Apicella CL, Marlowe FW, Fowler JH, Christakis NA. Social networks and cooperation in hunter-gatherers. Nature. Jan 26; 2012 481(7382):497–501. [PubMed: 22281599]

28. Rand DG, Arbesman S, Christakis NA. Dynamic social networks promote cooperation in experiments with humans. Proc Natl Acad Sci U S A. Nov 29; 2011 108(48):19193–19198. [PubMed: 22084103]

29. DesRoches CM, Campbell EG, Rao SR, et al. Electronic health records in ambulatory care--a national survey of physicians. N Engl J Med. Jul 3; 2008 359(1):50–60. [PubMed: 18565855]

30. Adams JL, Mehrotra A, Thomas JW, McGlynn EA. Physician cost profiling-reliability and risk of misclassification. N Engl J Med. 2010; 362(11):1014–1021. [PubMed: 20237347]

Technical Appendix

We used Episode Treatment Groups software to estimate the "intensity" of care delivered to patients cared for by each physician. In order to calculate costs that reflect differences in utilization rather than payment rates, we first calculated standardized costs for all Part A and Part B services received during the study period. Standardized cost differs from actual Medicare payment in two important ways. First, standardized cost incorporates the full allowed reimbursement from all payment sources (Medicare, patient cost sharing, and other insurers). Second, standardized cost eliminates the effects of various adjustments Medicare makes in setting local payment rates, such as geographic payment differences for local input price variations and differential payments across classes of providers (e.g. DSH and GME payments; cost-based reimbursement of critical access hospitals vs. DRG-based prospective payment for most other short-term hospitals). All costs were then adjusted to reflect CY 2006 reimbursement rates.

To calculate the intensity of care of an episode, wemultiplied the standardized cost for each service assigned to an episode by the number of times the service was delivered and summed the costs. We refer to this total as the observed cost. The observed costof an episode varies with the number of units of service delivered. We adjusted the data to eliminate extreme values for each episode type by truncating the charges to a minimum of $1/3^{rd}$ of the 25th percentile and a maximum of three times the 75th percentile. This differs slightly from other methods used in the literature, which typically entail truncating at a given percentile, but establishes clinically reasonable cut points.[30]

Episodes were attributed to the physician within each specialty who provided the most evaluation and management services for the care of that episode. We required that each "attributed" physician bill a minimum of 15% of the total evaluation and management costs for that episode. Consequently, multiple physicians of different specialties may be attributed to an episode of care. To calculate the intensity score, episodes were assigned to physicians who provided the most evaluation and management services within a specialty as long as this was greater than 30%.[30] These scores were then averaged across all episodes assigned to each physician in order to derive an intensity score for that physician. We did not assign ETG scores to physicians with fewer than 10 episodes in our care. We included a dummy variable in our model for dyads where one or both physicians were missing an ETG score.

Thresholding Analysis

As noted in the main text, we used an approach adapted from network science to apply a relative thresholding criterion for each node (physician). This allowed us to retain the strongest and most influential ties for each physician in the network, while also eliminating less important ties that might have arisen due to chance. As we noted, this method likely

eliminates some ties of importance, but maintains the basic topographical features of the network intact. To examine the sensitivity of this approach, we also explored different cut points using the top 10% of ties and the top 30% of ties. These results are presented in Appendix Table A1. As expected, measures such as the total number of included ties and adjusted degree per 100 patients treated are sensitive to this method (as they would be to any method of thresholding). As shown in the table, however, this method maintains the essential topological features of the network as shown by stability in the number included nodes, clustering, and PCP/Specialist relative centrality. There is also relative stability (e.g., between the top 30% v. the top 20% of ties) in network density.

Appendix Table A1

Thresholding Sensitivity Analysis

	No Threshold			30 Percent			20 Percent			10 Percent		
	Mean	**Min**	**Max**	**Mean**	**Min**	**Max**	**Mean**	**Min**	**Max**	**Mean**	**Min**	**Max**
Average Threshold	1.00	1.00	1.00	4.27	1.47	12.05	6.18	2.09	15.97	10.21	3.72	22.89
Number of nodes	1,339	135	8,197	1,339	135	8,197	1,339	135	8,197	1,337	135	8,197
Number of ties	108,634	4,103	695,672	70,743	2,134	549,393	50,927	1,568	392,582	28,827	830	226,194
Density	0.23	0.02	0.67	0.12	0.01	0.34	0.09	0.01	0.24	0.05	0.01	0.13
Adjusted Degree per 100 Patients	66.77	36.53	111.08	39.05	16.59	76.93	27.26	11.69	54.42	14.65	5.79	29.25
Clustering	0.63	0.42	0.85	0.56	0.38	0.70	0.55	0.40	0.67	0.53	0.44	0.63
PCP Relative Centrality	0.38	0.17	1.03	0.38	0.18	1.06	0.38	0.19	1.07	0.41	0.14	1.12
Specialists Relative Centrality	2.95	0.97	5.79	3.10	0.94	5.45	3.07	0.94	5.35	2.96	0.90	7.11

Figure A1.
US Map with HRRs highlighted

NUMBER OF DOC-DOC PAIRS

Panel a

NUMBER OF SHARED PATIENTS

NUMBER OF DOC-DOC PAIRS

Panel b

NUMBER OF SHARED PATIENTS
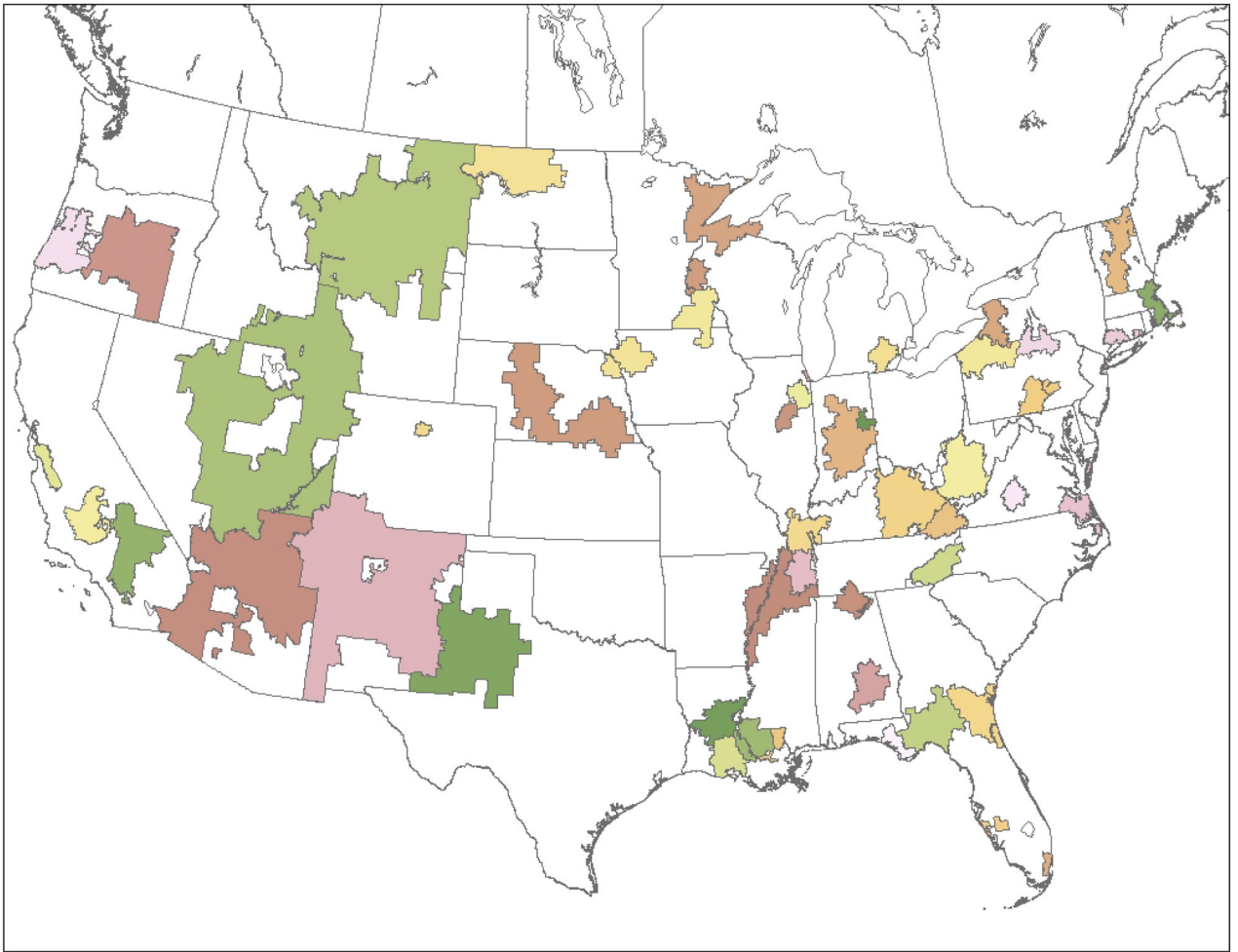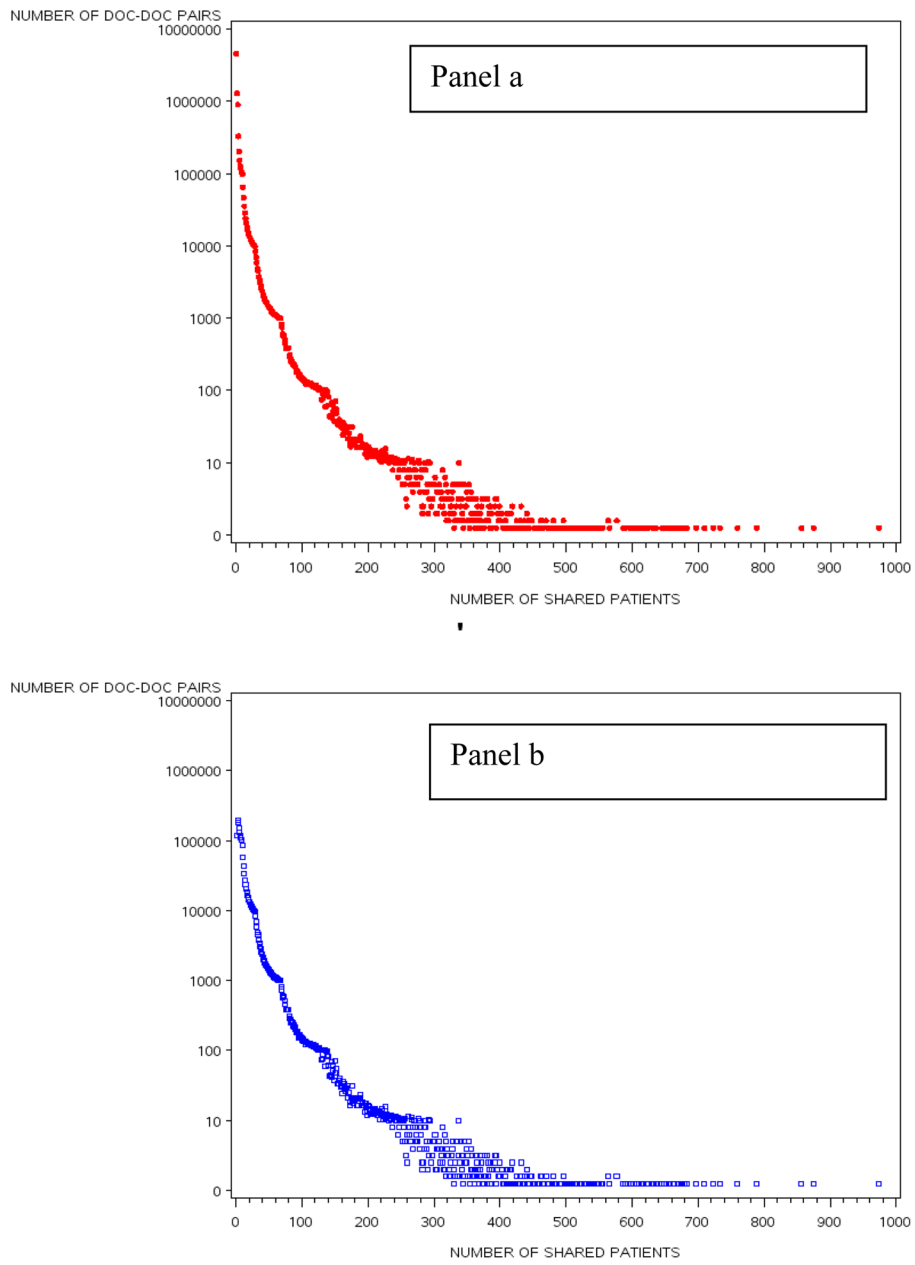
Figure A2.
Number of Shared patients by doctor pairs across the 51 HRRs Pre (Panel a) and Post (panel b) Thresholding

The number of shared patients is calculated after applying the top 20% thresholding rule. Adjusted degree (degree per 100 patients) is the average number of other physicians with whom a physician shares patients after adjusting for their total number of Medicare patients in 2006. The clustering coefficient is a metric that quantifies the extent to which the network neighbors of a given node are directly connected to one another. The *betweenness centrality* of a physician represents how central a physician is within his network of colleagues. To calculate relative PCP (or specialist) centrality, we divide mean PCP (specialist) centrality by the mean centrality of all other physicians in the network.
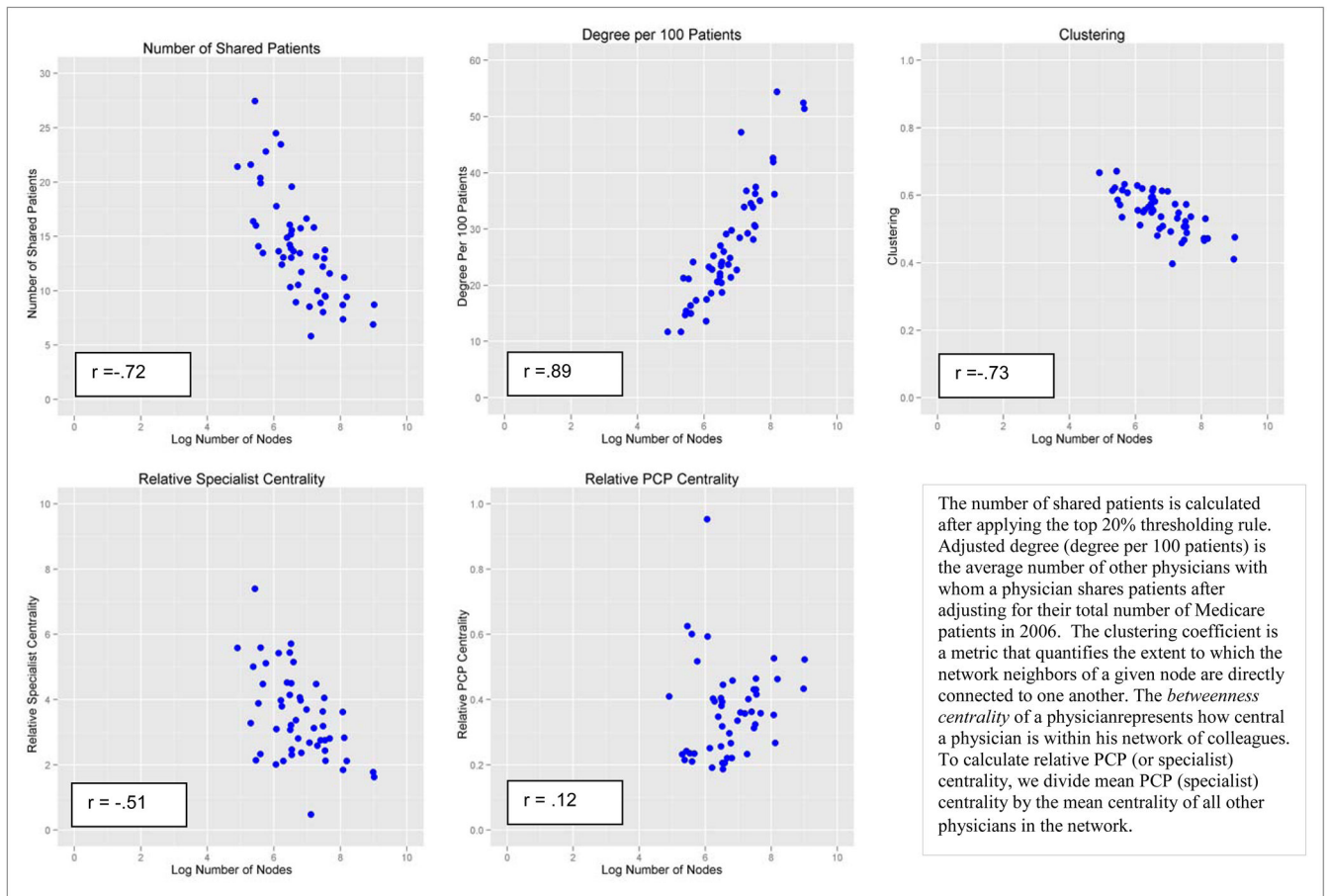
Figure A3.
Scatter Plots of Mean Network Attributes for 51 Hospital Referral Regions by Log Number of Physicians
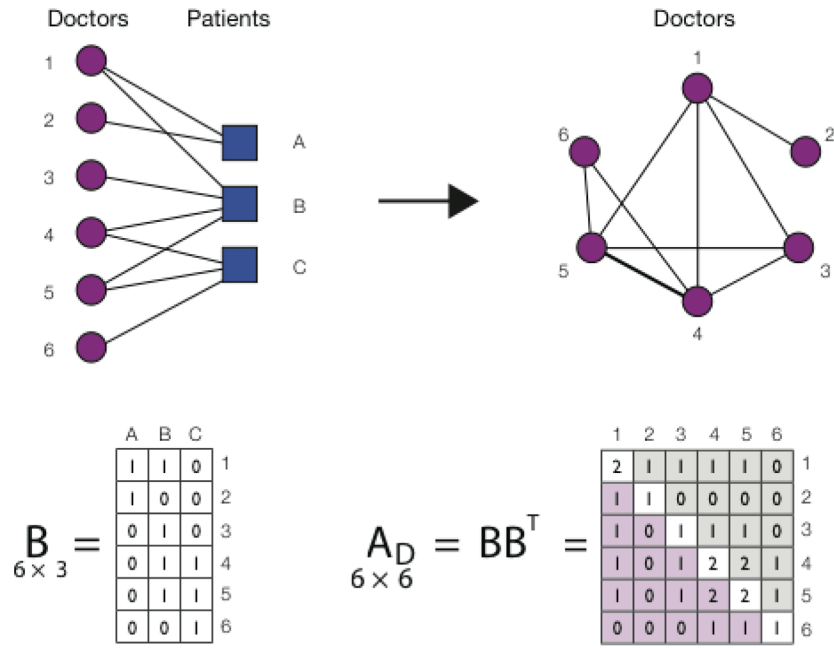
Figure 1.
A schematic illustrating a projection from a two-mode (bipartite) to a one-mode (unipartite) network. Medicare records link each doctor to a number of patients, which naturally leads to a bipartite network consisting of two types of nodes, doctors and patients. An edge can only exist between nodes of different type (a doctor and a patient), and the network is fully described by the (in this case 6×3) bipartite adjacency matrix B. A one-mode projection of the doctor-patient network is obtained by multiplying the bipartite adjacency matrix B by its transpose. The resulting symmetric one-mode adjacency matrix A is square in shape (in this case 6×6), and its elements indicate the number of patients the two physicians have in common. For example, $A(3,4) = 1$ shows that physicians 3 and 4 provide care for one common patient (patient B), whereas $A(4,5) = 2$ shows that physicians 4 and 5 have two patients in common (patients B and C). The diagonal elements of matrix A correspond to the number of patients the given physician provides care for, e.g. $A(4,4) = 2$ (in other words, physician 1 has degree 2).
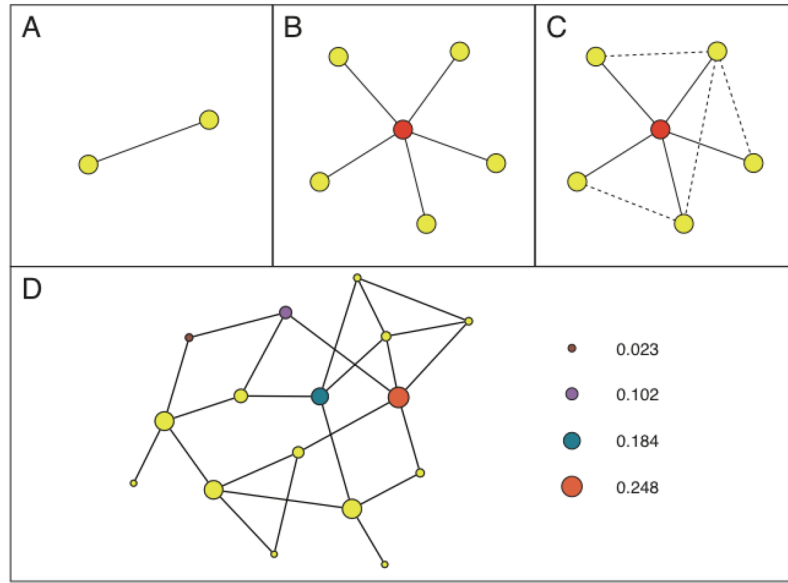
Figure 2.
This schematic illustrates some fundamental social networks concepts. (A) Nodes and ties are the elementary building blocks of networks. A tie connecting two nodes (physicians) indicates that the two physicians share one or more patients. (B) The *degree* of a node quantifies the number of connections a given node has. For example, the red node at the center of the figure has a degree of five, i.e. the physician shares patients with five other physicians. In our work, we present *adjusted degree* (degree divided by the number of Medicare patients cared for by each physician). (C) The *clustering* coefficient is a metric that quantifies the extent to which the network neighbors of a given node are directly connected to one another. More specifically, the clustering coefficient of the physician at the center of the figure (red node) is given by the number of ties that exist among his or her colleagues (the dashed four ties) divided by the number of ties that could exist between them (in this case 10), yielding a value of 4/10 or 0.4. This number can also be interpreted as the probability that any two randomly chosen network neighbors of the individual at the center are directly connected to one another. (D) *Betweenness centrality* quantifies the structural centrality of a node in the network. The betweenness centrality of a node is proportional to the number of times the node lies on shortest paths in the network, where one considers all shortest paths, i.e. shortest paths from every node in the network to every other node in the network. In the schematic, the size of a node is proportional to its betweenness centrality score, and the betweenness centrality scores are shown for four nodes in increasing order of centrality. In our work, we present relative PCP or specialist centrality, where we divide the mean PCP or specialist centrality for all physicians in the network by the mean centrality of all other physicians in the network.
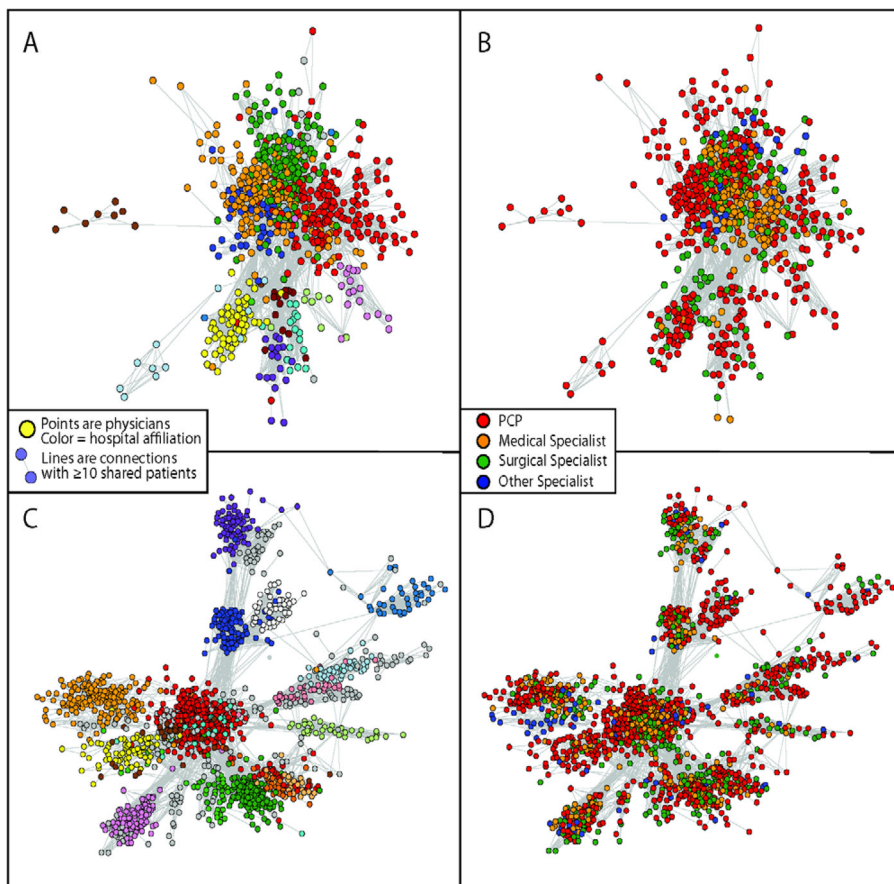
Figure 3.
Depictions of two networks: Albuquerque, NM (panels A and B, ~1000 physicians) and Minneapolis/St. Paul, MN (Panels C and D, ~1700 physicians). On the left (panels A and C), hospital affiliations are coded (each hospital is represented by a different color and on the right (panels B and D) specialty is coded.

Depictions of two networks: Albuquerque, NM (panels A and B, ~1000 physicians) and Minneapolis/St. Paul, MN (Panels C and D, ~1700 physicians) using "spring embedder" methods, which position objects with strong connections (i.e., physicians with more shared patients) in closer physical proximity. On the left (panels A and C), hospital affiliations are coded (each hospital is represented by a different color and on the right (panels B and D) specialty is coded.

Figure 4.
Scatter Plots of Mean Network Attributes for 51 Hospital Referral Regions by Log Number of Physicians (Nodes)

Figure 5.
Depictions of two networks: Albuquerque, NM (panels A and B, ~1000 physicians) and Minneapolis/St. Paul, MN (Panels C and D, ~1700 physicians). On the left (panels A and C), hospital affiliations are coded (each hospital is represented by a different color and on the right (panels B and D) specialty is coded.
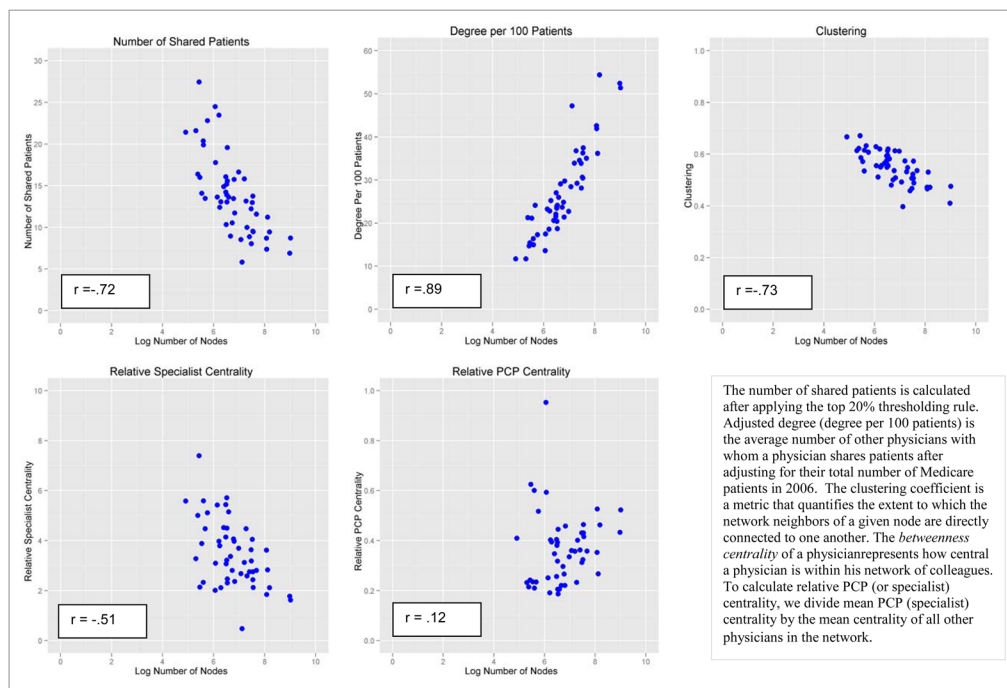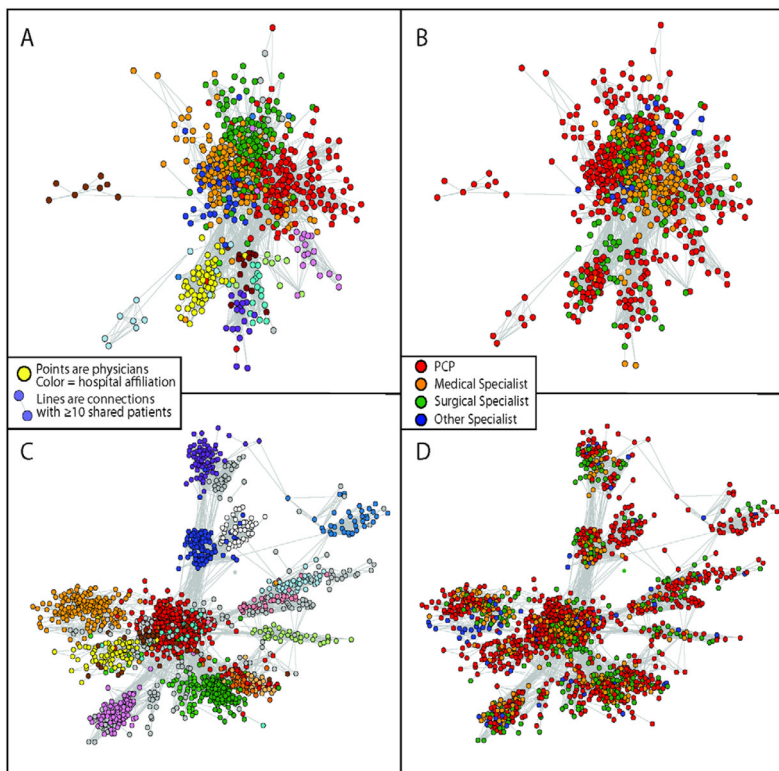
Table 1

Description of Physicians, Their Medicare Beneficiaries, and Network Characteristics stratified by HRR

	Mean (SD)	Range		Example HRRs			
		Min	Max	Boston, MA	Minot, ND	Joliet, IL	Miami, FL
Number of Physicians(nodes)	1,339	93	8,197	8,197	135	496	3,600
Physician Characteristics							
Mean Age	48.5 (1.3)	46.1	51.7	48.3	48.0	46.3	51.7
Percent Male	80.3 (4.9)	69.2	89.7	70.1	82.4	77.8	84.1
% Primary Care	41.9 (5.4)	27.6	53.3	38.5	50.4	40.1	39.0
% Medical Specialists	30.0 (4.8)	20.7	47.0	36.8	20.7	33.5	35.7
% Surgical Specialists	28.0 (3.3)	21.3	35.5	24.7	28.9	26.3	25.2
Patient Population Characteristics							
White (% of population)	86.6(11.5)	49.4	98.8	89.1	92.7	89.3	61.7
Black (% of population)	8.9 (10.1)	0.1	42.1	6.2	0.2	8.8	111.9
Hispanic (% of population)	1.8 (4.0)	0.0	24.2	1.5	0.0	1.0	24.2
Female (% of population)	59.9(2.0)	54.4	64.2	60.8	58.0	61.5	61.1
Medicaid(% of population)	23.0 (10.0)	7.7	51.9	27.7	12.0	14.8	51.9
Age (mean)	70.7 (1.6)	67.1	74.6	70.8	73.2	71.4	71.6
HCC Score (mean)	1.9 (.3)	1.5	2.8	2.2	1.7	2.0	2.8
ETG intensity Score (mean)	1.03 (.06)	0.9	1.2	1.1	1.2	1.0	1.1
# of patients (mean)	279 (84)	126	447	224	260	316	228
Network Characteristics							
Number of ties	50,927 (75,962)	1,568	392,582	392,582	1,568	12,914	218,136
Adjusted Degree	27.3 (10.4)	11.7	54.4	51.4	11.7	18.6	54.4
Number of Shared Patients	852 (336)	297	1,503	835	498	1222	1146
Clustering	.55 (.06)	.40	.67	.48	.67	.62	.47
PCP Relative Centrality	.38 (.17)	.19	1.06	.52	.41	.19	.46
Medical Specialist	3.47 (1.33)	.47	7.40	1.62	5.59	3.98	2.12

	Mean (SD)	Range		Example HRRs			
		Min	Max	Boston, MA	Minot, ND	Joliet, IL	Miami, FL
Relative Centrality							

HCC is hierarchical clinical conditions, ETG is Episode Treatment Group, HRR is hospital referral region. Thresholds refers to the mean number of shared patients required for establishing a tie in the respective HRRs. The SD is the standard deviation of the means.

Table 2

Physician and Patient Panel Characteristics Associated with Physician-Physician Relationships

	Unadjusted proportion of Dyads/Mean Difference with ties	Unadjusted proportion of Dyads/Mean Difference without ties	Unadjusted Rate Ratio[+]	Adjusted Rate Ratio[*+]
Physician Characteristics (Differences)	92.2	7.8	--	--
Sex				
Male-male	65.1	54.6	1.68 (1.68, 1.69)	1.32 (1.32, 1.32)
Female-Female	3.8	6.5	.72 (0.71,0.72)	.79 (.78, .79)
Male-Female	29.1	36.8	1.00 (ref)	1.00 (ref)
Difference in Age (mean) Specialty	11.5	12.5	.80 (.80, .80)	.88 (.88, .88)
PCP-PCP	10.1	15.8	.77 (.76, .77)	.62 (.62, .62)
PCP-Medical	28.0	27.1	1.00 (ref)	1.00 (ref)
PCP-Surgical	17.5	20.3	.72 (.72, .72)	.65 (.65, .65)
Medical-Medical	16.9	12.1	1.52 (1.52, 1.53)	1.36 (1.36, 1.36)
Medical-Surgical	20.7	17.9	.94 (.94, .95)	.90 (.89, .90)
Surgical-surgical	7.0	6.7	.76 (.76, .77)	.66 (.66, .66)
Distance (Mean)	13.2	24.2	.99 (.99, .99)	.98 (.98, .98)
Different Hospital (%)	69.2	96.0	.07 (.07, .07)	.12 (.12, .12)
Completed medical school at different medical school	6.1	3.6	.53 (.53, .53)	.99 (.99, .99)
Completed residency at different institution	5.3	2.7	.55 (.54, .55)	.88 (.88, .89)
Practice Style (ETG Intensity)	.29	.31	.91 (.91, .91)	.93 (.92, .93)
Patient Panel Characteristics[+]				
Difference in % White	11.5	20.2	.72 (.72, .72)	.89 (.89, .89)
Difference in % Black	8.8	14.0	.75 (.75, .75)	.92 (.92, .92)
Difference in % Hispanic	2.9	5.3	.59 (.59, .59)	.75 (.75, .76)
Difference in %Female	13.0	15.6	.80 (.80, .81)	.86 (.86, .86)
Difference in % Medicaid	15.3	24.4	.69 (.69, .69)	.86 (.86, .86)
Difference in Age (mean)	4.1	5.4	.42 (.42, .42)	.75 (.75, .75)
HCC Score (mean)	1.0	1.1	.78 (.78, .78)	.93 (.93, .93)

[*] Adjusted rate ratios and p-values were calculated using a negative binomial regression model, adjusting for all variables in the table. Rate ratios are used because the outcome (number of shared patients) is a count rather than binary. Results are similar when a binary outcome variable was analyzed using logistic regression. All p-values are less than .001

[+] Rate ratios reflect the increase in the expected number of shared patients (and thus likelihood of a true information sharing relationship) for every 10% point *difference* in patient panel characteristics (not applicable to HCC score). Ref signifies reference category

Glossary of Network Terms

Betweenness centrality	How central a node or a physician is within his network, obtained by considering shortest paths from each node to every other node in the network (Figure 2).
Bipartite network	"Bipartite" refers to a network where the nodes can be partitioned into two sets, here physicians and patients, such that all ties link nodes from one set to the other and there are no ties within a set. We convert this to a unipartite of linked physicians.
Clustering coefficient	The proportion of network neighbors of a node that are directly connected to one another, here a proportion of physician's colleagues who share patients with one another.
Connection, edge, or tie	A tie connects two nodes, in this case linking two physicians in the network who share patients as identified in Medicare claims data. Connections likely correspond to information sharing relationships between physicians.
Degree, adjusted degree	The number of ties a given node has, here the number of doctors a physician is connected to through patient sharing. Because patient volume influences the number of connections, we obtain adjusted degree by dividing degree by the total number of Medicare patients the physician shares with all other doctors.
Homophily	The tendency of individuals with similar characteristics to associate with one another.
Node or actor	An individual actor or agent in the network, in this case a physician.
Relative betweenness centrality	The mean betweenness centrality of one physician type (e.g., PCP) relative to all other physicians in the network.
Shared patients	The total number of shared patients across *all* ties for an individual physician.
Social network	A set of actors, in this case physicians, and a set of relationships linking the actors together. Social networks can be used to study the structure of a social organization and how this structure influences the behavior of individual actors.