A measure of local uniqueness to identify linchpins in a social network with node attributes

Matthew D. Nemesure[1†], Thomas M. Schwedhelm[1†], Sofia Sacerdote[2], A. James O'Malley[1,3], Luke R. Rozema[1] and Erika L. Moen[1,3*]

*Correspondence:
Erika.L.Moen@dartmouth.edu
†Co-first author: Matthew D. Nemesure and Thomas M. Schwedhelm
[1] Department of Biomedical Data Science, Geisel School of Medicine at Dartmouth, One Medical Center Dr, Lebanon, NH 03756, USA
Full list of author information is available at the end of the article

Abstract

Network centrality measures assign importance to influential or key nodes in a network based on the topological structure of the underlying adjacency matrix. In this work, we define the importance of a node in a network as being dependent on whether it is the only one of its kind among its neighbors' ties. We introduce *linchpin score*, a measure of local uniqueness used to identify important nodes by assessing both network structure and a node attribute. We explore linchpin score by attribute type and examine relationships between linchpin score and other established network centrality measures (degree, betweenness, closeness, and eigenvector centrality). To assess the utility of this measure in a real-world application, we measured the linchpin score of physicians in patient-sharing networks to identify and characterize important physicians based on being locally unique for their specialty. We hypothesized that linchpin score would identify indispensable physicians who would not be easily replaced by another physician of their specialty type if they were to be removed from the network. We explored differences in rural and urban physicians by linchpin score compared with other network centrality measures in patient-sharing networks representing the 306 hospital referral regions in the United States. We show that linchpin score is uniquely able to make the distinction that rural specialists, but not rural general practitioners, are indispensable for rural patient care. Linchpin score reveals a novel aspect of network importance that can provide important insight into the vulnerability of health care provider networks. More broadly, applications of linchpin score may be relevant for the analysis of social networks where interdisciplinary collaboration is important.

Keywords: Node attribute, Patient-sharing networks, Network vulnerability, Centrality

Introduction

Centrality is one of the most established concepts in social network analysis. One of its prime applications is to identify important or influential nodes based on their structural position. Centrality measures such as degree, betweenness, and closeness centrality were introduced by Freeman in the 1970s (Freeman [1977], [1978]). Eigenvector centrality, another measure of influence, was introduced by Bonacich and identifies influential nodes based on the prominence of their direct ties (Bonacich [1972]). Each centrality

measure was developed to infer distinct theorized aspects of importance or influence based on the topological characteristics of the adjacency matrix of the network in which a node is embedded.

More recently, advancements in network centrality have included community-aware centrality and multi-component centrality measures. Community-aware centrality measures identify nodes that are essential to connect two or more communities of the network (Tulu et al. 2018). Extensions from this work have defined influence based on the extent to which a node is a hub within their community and a bridge across communities (Ghalmane et al. 2019a, b). Redefining local and global influence in networks with overlapping communities, new representations of centrality measures have been developed that are specifically designed to identify influential nodes in overlapping modular networks (Ghalmane et al. 2019a, b).

These structural centrality measures remain agnostic to node attributes. Node attributes are used to describe characteristics and can be continuous or discrete. Widely used network measures that do consider node attributes include assortativity and homophily, which are network-level measures of the correlation (assortativity) or tendency (homophily) of nodes to be connected to similar others (Newman 2003; McPherson et al. 2001). Many studies have observed assortativity or homophily in social networks by characteristics such as happiness, smoking and drinking behavior, and race (Bollen et al. 2011; Bliss et al. 2012; Cheadle et al. 2013; Smith et al. 2014; Mollica et al. 2003). Community detection is another example of an established network algorithm that has evolved to consider node attributes (Zanghi et al. 2010; Newman and Clauset 2016; Jia et al. 2017). Network communities can be identified by combining structural and attribute information such that communities consist of nodes that are not only more densely connected than nodes outside of the community, but also share similar attributes (Jia et al. 2017).

This collection of work provides strong evidence that the attributes of individuals in a network relate to, and even influence, network structure. This raises the question of how attributes can be leveraged to identify strategically important nodes in a way that is distinct from the centrality measures that rely purely on the underlying adjacency matrix. In line with this question, previous work has decomposed centrality measures according to categorical attribute data (Everett and Borgatti 2012; Krackhardt and Stern 1988; Moen et al. 2019). The present work introduces a new node-level measure that combines the topological data from the adjacency matrix with accompanying external attribute data. Herein, we propose *linchpin score*, which describes the tendency of a node to be the only one of its kind among its neighbors' ties. The term linchpin was chosen because it refers to nodes that are indispensable within its two-hop neighborhood. We consider a node to be more indispensable, or a linchpin, if more of its neighbors have no other existing ties to other similar nodes. This term thus defines the importance of a node in a network as being dependent on whether or not its neighbors are directly connected to others that are similar to the focal node.

Motivation

Traditional health services research methods evaluate the quality of individual physicians as a function of other physician-level attributes. However, physicians are embedded within professional networks, and their individual outcomes and ability to deliver

Nemesure *et al. Appl Netw Sci*   (2021) 6:56

Page 3 of 14

high quality care may be impacted by their own position in their peer network or the characteristics and outcomes of their peers. Patient-sharing networks offer a quantitative, scalable approach for indirectly measuring relationships between physicians based on shared patients observed in administrative data. Prior work has shown that patient-sharing relationships correspond with self-reported referral and advice-seeking relationships between physicians (Barnett et al. 2011). Patient-sharing network characteristics have been associated with care utilization, care quality, and patient outcomes (Pollack et al. 2013; Bachand et al. 2018; Tannenbaum et al. 2018; Moen et al. 2016; 2019; Zipkin et al. 2021; Barnett et al. 2012). Increased patient-sharing within physician group practices has been shown to correspond with patient-reported care coordination and timeliness of care (Moen and Bynum 2019). However, while increased patient-sharing within teams of physicians is hypothesized to reflect care coordination, the absence of ties to other physicians may suggest barriers in access to specialist referrals or other important resources (Hollingsworth et al. 2015).

Health care delivery systems depend on the availability of personnel and infrastructure to deliver high quality care. The shortage of medical professionals in rural areas is a significant national concern. Access to health care is typically measured according to the supply per capita or distance to one type of provider or service (Levit et al. 2020). The National Rural Health Association reports 13.1 physicians per 10,000 people in rural areas compared with 31.2 physicians per 10,000 people in urban areas. The number of specialists per capita is even more skewed, with 30 specialists per 10,000 people in rural areas compared with 263 specialists per 10,000 people in urban areas. Given the importance of multidisciplinary care coordination in the delivery of high-quality care for many complex and chronic conditions, there is a strong premise for using networks to understand access to care which recognizes the importance of professional relationships.

In creating this measure, we also take inspiration from the concept of network vulnerability to selective node removals (Chen and Hero 2013). One of the most conventional network vulnerability measures is the susceptibility of the size of the largest connected component to the removal of nodes. In this sense, a node would be considered more vital to the network if the largest connected component was more disrupted (e.g., broken into smaller, disconnected components) upon its removal. Many studies assessing network vulnerability focus on infrastructure networks (Grubesic and Murray 2006; Corley and Chang 1974). Yet social networks can also be vulnerable to disruption upon an individual node's removal. Here, we hypothesize that the neighborhood of a physician would be more vulnerable to the physician's removal if their neighbors have no existing connections to other physicians of the same specialty. While this measure has implications to a broader range of network studies, an application of linchpin score in health services research would be to identify networks or sub-networks that are more vulnerable to removal of physicians of a specific specialty.

The rest of the paper is structured as follows. We next propose and formally define linchpin score. Then, we measure linchpin score in a physician network using specialty as the node attribute of interest. We calculate linchpin score for the physicians in the network, summarize linchpin score by specialty, and compare the observed linchpin scores to those measured in random networks. We then evaluate whether linchpin score is associated with degree, betweenness, closeness, and eigenvector
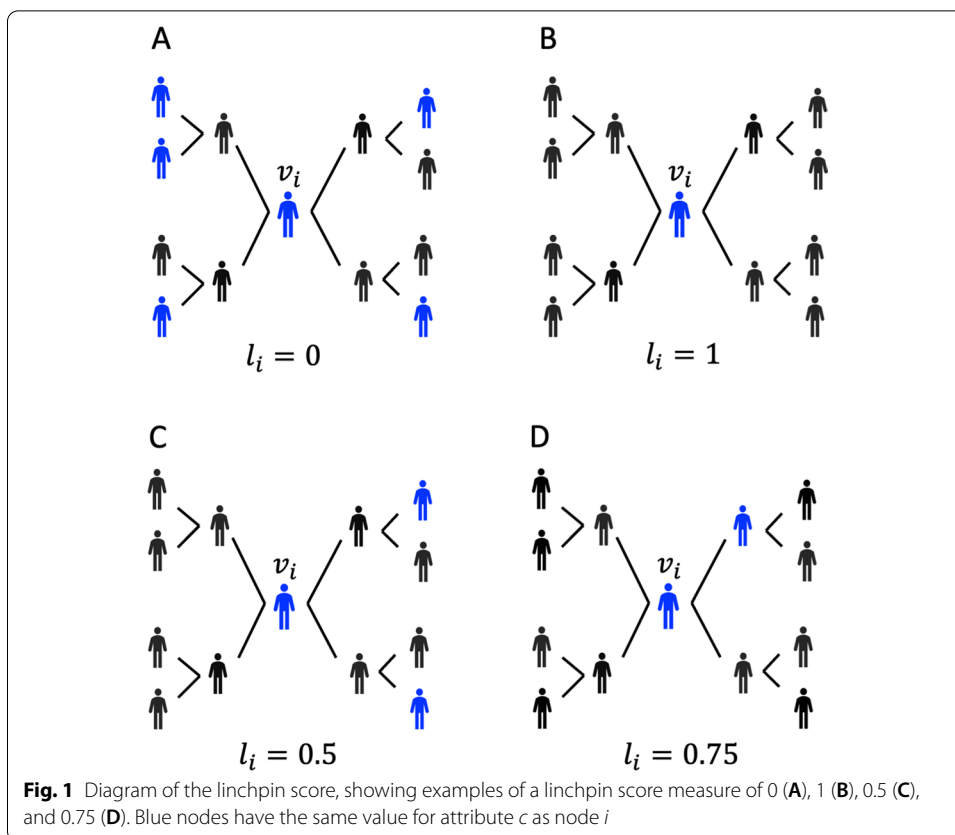
centrality. Finally, we examine linchpin score within 306 patient-sharing physician networks, representing the 306 hospital referral regions in the United States. We test the extent to which linchpin score and other centrality measures are associated with physician rurality within hospital referral regions and across specialty types.

Methods

A graph $G = (V, E)$ consists of a set of nodes V and a set of edges E between them. An edge e_{ij} connects node v_i with node v_j. Let c_i denote the type of node v_i for attribute c, where $i = 1, \ldots, N$ is the index of the nodes in the network. While the applications in this work focus on categorical node attributes, linchpin score can be extended to continuous node attributes by setting a threshold to bin the continuous variable into discrete categories. The linchpin score for node v_i, denoted by l_i, is the number of neighbors of node v_i with no other ties to any other node equal to node v_i for attribute c, divided by n_i, the degree of node v_i. The neighbors of node v_i are not allowed to have the same attribute value as node v_i to contribute to l_i. Let the event that nodes i and j have the same value of attribute c be denoted by the binary variable a_{ij}^c. That is, $a_{ij}^c = 1$ if $c_i = c_j$ and $a_{ij}^c = 0$ otherwise. The number of neighbors v_k of node v_j for which $c_k = c_i$ other than node i itself is given by $d_{ij}^c = \sum\limits_{k \neq i,j}^{N} e_{jk} a_{ik}^c$. Defining $b_{ij}^c = 1$ if $d_{ij}^c > 0$ and $b_{ij}^c = 0$ otherwise, the mathematical definition of linchpin score is then expressed as:

$$l_i = \frac{1}{n_i} \sum_{\substack{j \neq i}}^{N} e_{ij} \left(1 - a_{ij}^c\right) \left(1 - b_{ij}^c\right)$$

The linchpin score of node i is seen to be the weighted degree of node i, where the weights for node $j \neq i$ are given by $\left(1 - a_{ij}^c\right)\left(1 - b_{ij}^c\right)$, divided by the degree of node i. The first term comprising the weight, $1 - a_{ij}^c$, indicates whether nodes i and j have different values of attribute c while the second term, $1 - b_{ij}^c$, indicates whether none of the other neighbors of node j (besides node i) have the same value of attribute c as node i. The linchpin score ranges from 0 to 1, with 0 indicating all of the neighbors of node v_i are connected to at least one node that is equal to node v_i for attribute c (Fig. 1A), and 1 indicating that none of the neighbors of node v_i are equal to node v_i for attribute c nor are connected to another node (besides node i) that is equal to node v_i for attribute c (Fig. 1B). Figure 1C illustrates the linchpin score for a node v_i who has two out of four connections also tied to another node that is equal to node v_i for attribute c. In Fig. 1D, we consider the circumstance where node v_i is directly connected to another node of the same value for attribute c. In this case, we do not count the neighbor with the same attribute value as node v_i in the calculation of l_i, but that neighbor would still contribute to n_i. We take this approach because it would be reasonable to expect that the focal node's direct ties could relatively easily form a new tie with the neighbor that has the same attribute value as the focal node, if the focal node were to be removed from the network. The R code to calculate linchpin score on any network dataset that contains node attributes is available at https://github.com/mnemesure/linchpin_centrality/blob/master/linchpin_network_fx2.R.

Nemesure *et al. Appl Netw Sci*   (2021) 6:56

Page 5 of 14

Fig. 1 Diagram of the linchpin score, showing examples of a linchpin score measure of 0 (**A**), 1 (**B**), 0.5 (**C**), and 0.75 (**D**). Blue nodes have the same value for attribute *c* as node *i*
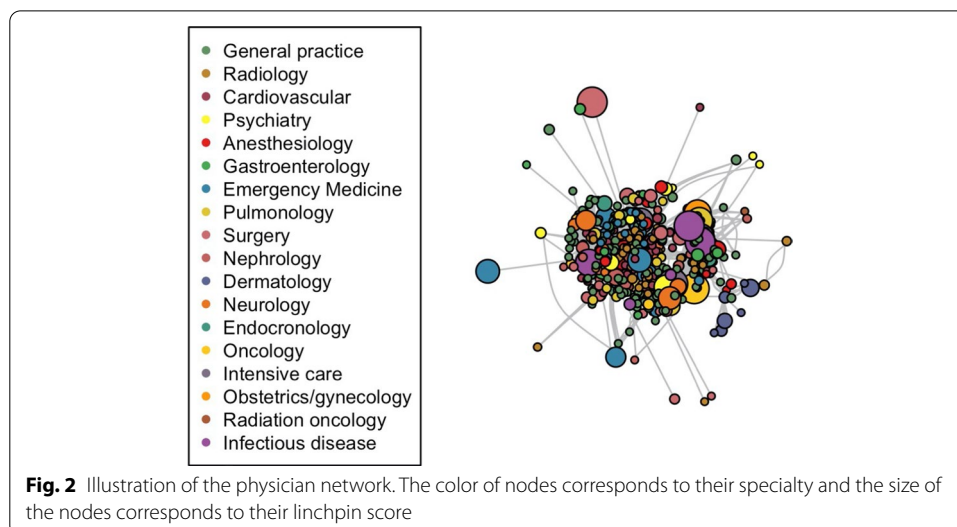
Example network datasets

For the patient-sharing network analysis, we linked four publicly available data sources. The first data source is the Physician Shared Patient Patterns Data from 2015 released by Centers for Medicare and Medicaid Services (CMS) (Physician Shared Patient Patterns 2015). The Physician Shared Patient Patterns Data lists health care physicians who participate in the delivery of health services to the same Medicare beneficiary within specific time intervals (30 days, 60 days, 90 days, and 180 days). It reports the number of patients each physician dyad shared within the specified time interval. We used the Physician Shared Patient data to create undirected patient-sharing networks for which ties between physicians indicate shared patients within 30 days in 2015.

The second data source was the November 2015 Physician Compare National Downloadable File released by CMS and archived by the National Bureau of Economics Research (Physician Compare 2015). The dataset contains general information about individual eligible health care professionals including specialty, practice affiliation, and practice ZIP code. For the purposes of this study, we used this dataset to obtain specialty and practice ZIP code.

The third data source links ZIP codes to hospital referral regions as defined and made available by the Dartmouth Atlas (Dartmouth Atlas Supplemental Data 2015). Hospital referral regions represent regional health care markets for tertiary medical care, and there are 306 geographically contiguous hospital referral regions in the US. We assigned each physician to a hospital referral region based on their practice ZIP code. This further

Fig. 2 Illustration of the physician network. The color of nodes corresponds to their specialty and the size of the nodes corresponds to their linchpin score

allowed us to parse the national network into sub-networks that represent the patient-sharing patterns within regional health care delivery markets. We examined the Providence, RI hospital referral region as our example network due to its relatively small size and our familiarity with the area.

The fourth dataset includes the 2010 Rural–Urban Commuting Area (RUCA) codes for each ZIP code released by the United States Department of Agriculture Economic Research Service updated on August 17, 2020 (Rural–Urban Commuting Area Codes 2020). Rural physicians were identified as those who practice in a rural ZIP code (RUCA codes 4.0-10.6) based on the practice ZIP codes listed in Physician Compare National Downloadable File. Physicians who practice in multiple locations that included both rural and urban ZIP codes were categorized as urban.

Results

Linchpin scores of physicians in a patient-sharing network

Networks of physicians are frequently assembled based on administrative data of patient encounters: two physicians are connected if they have encounters with common patients. In this example, we evaluate linchpin score of physicians using physician specialty as the node attribute of interest. If a physician is the only one of their specialty among their neighbor's ties, it is reasonable to expect that the physician is indispensable for the proper coordination and delivery of health care to the patients cared for by that set of physicians.

Primary specialty data was simplified into 18 specialty categories: cardiovascular, surgery, obstetrics/gynecology, general practice, psychiatry, infectious disease, pulmonology, endocrinology, gastroenterology, neurology, anesthesiology, radiology, emergency medicine, nephrology, oncology, dermatology, radiation oncology, and intensive care.

The Providence, RI physician network includes 1,749 physicians, has a density of 0.017 and global transitivity of 0.296 (Fig. 2). The mean linchpin score by specialty type varies substantially (Table 1). Intensive care, obstetrics-gynecology, and endocrinology are the specialties with the highest mean linchpin score (0.88, 0.43, and 0.40, respectively), whereas radiologists and cardiologists had the lowest mean linchpin scores. We evaluated the correlation between the number of physicians in the specialty category

Nemesure *et al. Appl Netw Sci*   (2021) 6:56

Page 7 of 14

Table 1 Summary statistics of linchpin score by specialty type

	N	Linchpin mean (SD) Observed network	Linchpin mean (SD) Random networks
Physician network attribute			
Anesthesiology	92	0.09 (0.19)	0.19 (0.28)
Cardiovascular	93	0.02 (0.04)	0.18 (0.28)
Dermatology	41	0.32 (0.32)	0.28 (0.30)
Emergency medicine	91	0.08 (0.19)	0.15 (0.25)
Endocrinology	31	0.40 (0.36)	0.32 (0.31)
Gastroenterology	54	0.14 (0.23)	0.24 (0.29)
General practice	668	0.06 (0.19)	0.04 (0.15)
Infectious disease	22	0.25 (0.34)	0.42 (0.34)
Intensive care	4	0.88 (0.12)	0.77 (0.22)
Nephrology	26	0.08 (0.12)	0.40 (0.34)
Neurology	50	0.21 (0.28)	0.27 (0.31)
Obstetrics/gynecology	26	0.43 (0.45)	0.36 (0.32)
Oncology	49	0.16 (0.26)	0.27 (0.31)
Psychiatry	87	0.16 (0.31)	0.18 (0.28)
Pulmonology	33	0.13 (0.20)	0.31 (0.31)
Radiation oncology	14	0.31 (0.32)	0.57 (0.34)
Radiology	118	0.02 (0.10)	0.15 (0.25)
Surgery	250	0.09 (0.19)	0.09 (0.21)

The number of nodes (N) represents the number of physicians with that specialty in the Providence, RI physician network. The linchpin mean and SD for the observed network are presented alongside the linchpin mean and SD for 10 permuted networks where the network structure was constant but specialty was randomly assigned. SD, standard deviation

and mean linchpin score using Kendall's Tau, a non-parametric correlation coefficient. Examining the mean linchpin score by specialty type reveals an inverse association with the number of physicians who have that specialty (Kendall's $\tau = -0.6$, $p < 0.001$). In other words, specialties that are rarer tend to have higher linchpin score.

By comparing linchpin score of physicians within the same specialty, one would identify physicians who are more indispensable to ensure that other members of the network have access to that specialty for referrals. Those physicians with greater linchpin score would be less easily replaced by another physician of their same specialty based on existing ties if they were to leave the network. Specialties with the highest variance in linchpin score among physicians of that specialty type are obstetrics-gynecology, infectious disease, and endocrinology (Table 1).

Next, we compared the linchpin scores of specialties in the observed physician network to a network in which specialty is distributed at random. We permuted ten networks that were identical to the observed network in structure and number of nodes with each specialty, but specialty was assigned at random. We then calculated the mean and standard deviation of linchpin scores of physicians in each specialty and present the means across the 10 permuted networks. The negative association between mean linchpin score and number of physicians who have that specialty was even stronger in the random networks (Kendall's rank correlation $\tau = -0.9$, $p < 0.001$) than what we found for the observed network. We also found that 12 of the 18 specialty groups have lower mean linchpin score in the observed network compared with the random networks (Table 1).

Fig. 3 Heatmap of the correlation matrix of linchpin, degree, betweenness, closeness, and eigenvector centrality for the physician network. Correlation was measured using Kendal's Tau non-parametric correlation coefficient
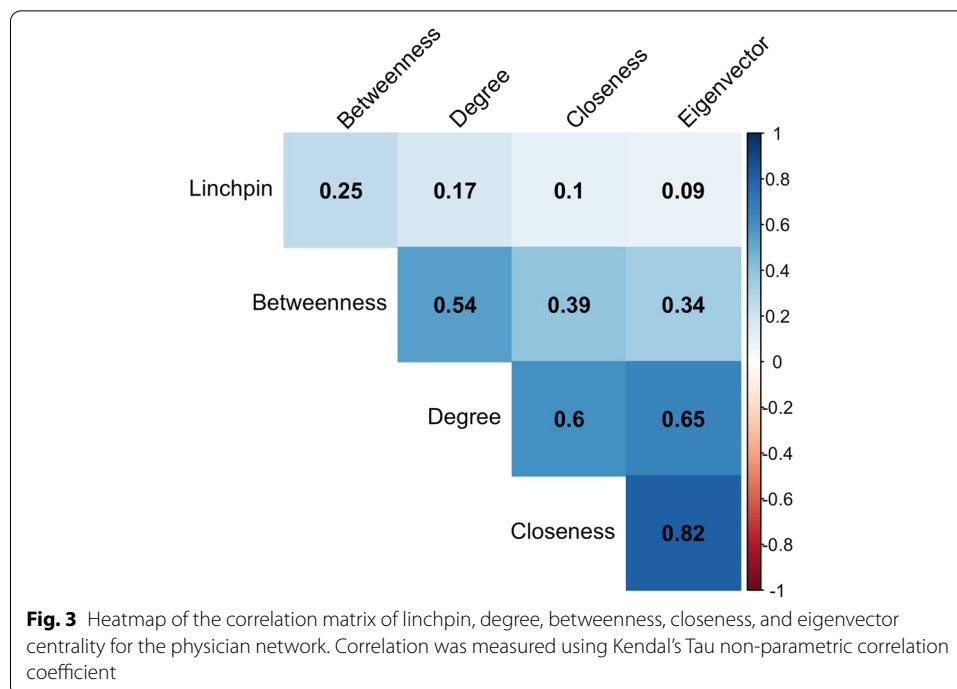
Taken together, these results suggest that the patient-sharing patterns may have formed in ways that make the networks less vulnerable, or less dependent on linchpin physicians.

Correlations between linchpin score and centrality measures

To examine whether a physician's linchpin score was associated with node centrality measures, we present a correlation matrix for linchpin, degree, betweenness, closeness, and eigenvector centrality for the Providence, RI physician network (Fig. 3). Previous work has demonstrated that network centrality measures tend to be correlated (Valente et al. 2008; Rajeh et al. 2020). We find that linchpin score is modestly correlated with betweenness centrality (Kendall's $\tau = 0.25$, $p < 0.001$). In general, linchpin score seems to be identifying a distinct set of important nodes in each network that are not captured by the other centrality measures and *vice-versa*. Consistent with previous work, we observed moderate to high correlations between the node centrality measures. Closeness centrality and eigenvector centrality were most highly correlated (Kendall's $\tau = 0.82$, $p < 0.001$).

Linchpin score and physician rurality

The motivation for developing linchpin score was to identify locally unique physicians who would not be easily replaced by another physician of the same specialty through existing ties if they were to leave the network. Linchpin score is most relevant for attributes that are difficult to change. For example, a physician cannot easily change specialties. Networks characterized by high linchpin score for a specialty of interest could be considered more vulnerable to the removal of physicians with that specialty. We aimed

to test this with a study of physicians practicing in rural and urban areas in the United States.

As a consequence of the uneven distribution of specialists across rural and urban areas, we expect that physician networks caring for predominantly rural patients may differ in both the organization of ties and the types of specialty groups present compared with physician networks caring for predominantly urban patients. We first examined associations between physician rurality and the node-level network measures. Then, we examined each specialty separately to determine whether the node-level measures were able to distinguish differences in network importance among rural and urban physicians by specialty type. We consider rural areas as being more vulnerable to a specialist leaving and we hypothesized a priori that rural specialists will have higher linchpin scores compared with urban specialists. We further hypothesized that this association may not be observed for general practitioners, who are more prominent in the care of rural patients.

We calculated the linchpin score, degree, betweenness, closeness, and eigenvector centrality for all physicians within all 306 hospital referral region networks. All network measures were standardized to have a mean of 0 and a standard deviation of 1 to better compare model estimates between network measures across hospital referral regions of different sizes. Eigenvector centrality was excluded from the models due to issues of high collinearity. We first examined associations between physician rurality and network measures within hospital referral regions. Physician rurality was represented as a binary variable assigned based on the rurality of the practice ZIP code, as defined in the methods. We excluded regions where fewer than 3% of physicians practiced in a rural ZIP code (n = 147), as some hospital referral regions are entirely urban. For each of the 159 hospital referral regions remaining, we estimated a separate multivariable logistic regression predicting physician rurality. Physician linchpin score, degree, betweenness, and closeness centrality were the independent variables of interested and we included physician specialty as a covariate. Based on the model results, we calculated the number of hospital referral regions for which each network measure was a significant predictor of rurality (corresponding to a *p*-value less than 0.01) and the number of hospital referral regions for which each network measure was the strongest predictor of rurality (corresponding to the highest z value). Our results, shown in Table 2, demonstrate that closeness centrality is more likely to be associated with physician rurality compared with the other network measures. This suggests that the differences in physician connectedness between rural and urban physicians is best detected using a centrality measure based on average distances to other nodes. This may reflect the regionalization of health care, often embodied by a regional "hub" with spokes extended to adjacent, more rural settings.

Table 2 Number of hospital referral regions for which each network measure significantly predicts physician rurality
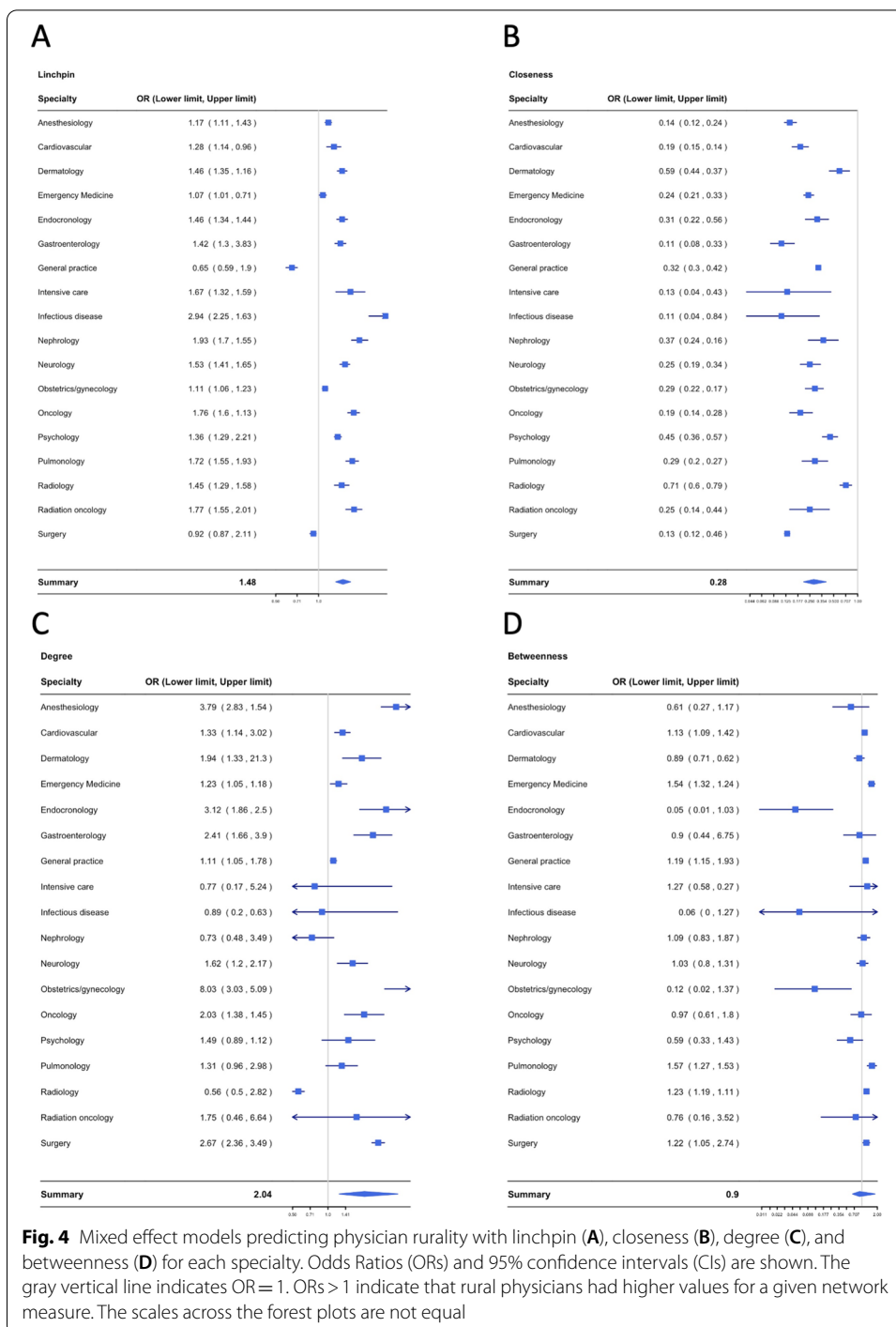
Network measure	Significant predictor of rurality (*p* < 0.01) n (%)	Strongest predictor of rurality (greatest absolute z value) n (%)
Linchpin	45 (28)	17 (11)
Degree	55 (35)	29 (18)
Closeness	96 (60)	100 (63)
Betweenness	41 (26)	13 (8)

To learn more about the characteristics of hospital referral regions where closeness centrality or linchpin score was predictive of physician rurality, we further characterized these hospital referral regions using network measures such as network size (e.g., number of physicians), network density, and network transitivity. Network density and transitivity have previously been shown to impact the relationships among node-level measures, such as the correlation between centrality and hierarchy measures (Rajeh et al. 2020). We also evaluated the proportion of physicians within each hospital referral region who practiced in a rural setting. With bivariate analyses, we found that hospital referral regions where closeness centrality was a significant predictor of rurality were characterized by networks of larger size ($p < 0.001$), lower density ($p < 0.001$), and lower transitivity ($p < 0.001$). Hospital referral regions where linchpin score was a significant predictor of rurality were also characterized by networks of larger size ($p = 0.02$), lower density ($p < 0.001$), and lower transitivity ($p = 0.01$). The performances of closeness centrality and linchpin score were not associated with the proportion of rural physicians.

Next, we evaluated associations between network characteristics and rurality within each of the 18 specialty groups. For each specialty group, using data from all 306 hospital referral region networks, we developed separate mixed effect logistic regressions predicting rurality of physician with linchpin score, degree, betweenness, and closeness centrality as independent variables. We included network size, network density, and network transitivity as covariates, and included a random effect for hospital referral region.

In Fig. 4, we show the adjusted associations between physician rurality and each of the network measures. To facilitate comparisons in the association between centrality and rurality across specialty types, we grouped the results by network measure. We observe significantly greater linchpin scores for rural physicians across almost all specialty groups, highlighting the importance of individual specialists in delivering services specific to that specialty among their direct ties. The only specialty groups that exhibit lower linchpin score in rural areas are general practitioners and, to a lesser effect, surgeons. These results are consistent with our hypothesis and provide additional evidence that general practitioners are more prominent in managing care for rural patients. They are more likely to be either directly connected or indirectly connected to each other through referrals to other specialists in their local networks, resulting in lower linchpin score. The patterns across all specialty groups for the other centrality measures varied. Notably, linchpin score is the only measure that distinguishes rural specialists from rural general practitioners. Closeness centrality was consistently lower across all specialty types, including general practitioners, in rural areas compared with urban. Degree centrality of rural physicians compared with urban tended to be either greater or not significantly different. Betweenness centrality did not show a strong association with the rurality of physicians across most specialties.

Altogether, these results demonstrate that incorporating specialty in defining physician network characteristics adds important contextual information to understanding which physicians are important. Closeness centrality was lower among all rural physician specialty groups compared with their urban counterparts, indicating that while closeness centrality was a strong predictor of rurality, it was not able to pick up differences between types of physicians in terms of which were locally unique for rural patients, which is an important aspect of rural health care access and quality. Linchpin

Fig. 4 Mixed effect models predicting physician rurality with linchpin (**A**), closeness (**B**), degree (**C**), and betweenness (**D**) for each specialty. Odds Ratios (ORs) and 95% confidence intervals (CIs) are shown. The gray vertical line indicates OR = 1. ORs > 1 indicate that rural physicians had higher values for a given network measure. The scales across the forest plots are not equal

score, on the other hand, did not tend to be the strongest predictor of rurality, but it was able to distinguish the different roles in the network played by specialists and general practitioners.

Conclusions

In this work, we have introduced a novel node-level measure, linchpin score, which identifies important nodes based on the extent to which they are the only one of their kind among their neighbors' connections. Other recent advancements in network centrality have developed measures of importance that combine information on the position of a node in a network with the local information on its nearest neighbors (Ibnoulouafi et al. 2018). Unlike centrality measures that are based solely on the topological structure of the network, linchpin score identifies a distinct set of locally unique nodes in a network using attribute data.

Our motivation for developing linchpin score was to identify physicians who were linchpins in their local networks. We evaluated linchpin score by specialty in 306 patient-sharing physician networks to characterize physicians who were the only physician of their specialty type among their neighbors' ties. The results from our regression models support our hypothesis that high linchpin score is associated with rurality of specialist physicians, indicating the vital role specialists play in delivering care to rural patient populations. We anticipate that linchpin score will offer a network-based metric of indispensable physicians in a health care delivery market. Similar to vertex coloring in graph theory (Galinier et al. 2013), networks could be designed so that the attributes (or colors) of the vertices are arranged to optimize a measure (e.g., linchpin score) in a graph, subject to certain constraints. Health systems aiming to reduce the vulnerability of their physician networks could use network simulations to identify more robust networks that reduce linchpin scores of specialties of interest. While the motivation for this work was to address a limitation in our ability to identify important physicians in local networks, linchpin score can be calculated for any network with node attributes, extending its applicability across social network research domains.

Nemesure *et al. Appl Netw Sci*   (2021) 6:56

Page 13 of 14

Declarations

Author details
[1]Department of Biomedical Data Science, Geisel School of Medicine at Dartmouth, One Medical Center Dr, Lebanon, NH 03756, USA. [2]Brown University, Providence, RI 02912, USA. [3]The Dartmouth Institute for Health Policy and Clinical Practice, Geisel School of Medicine at Dartmouth, Lebanon, NH 03756, USA.

References

Bachand J, Soulos PR, Herrin J, Pollack CE, Xiao Xu, Ma X, Gross CP (2018) Physician peer group characteristics and timeliness of breast cancer surgery. Breast Cancer Res Treat 170(3):657–665. https://doi.org/10.1007/s10549-018-4789-8

Barnett ML, Landon BE, O'Malley AJ, Keating NL, Christakis NA (2011) Mapping physician networks with self-reported and administrative data. Health Serv Res 46(5):1592–1609. https://doi.org/10.1111/j.1475-6773.2011.01262.x

Barnett ML, Christakis NA, O'Malley J, Onnela J-P, Keating NL, Landon BE (2012) Physician patient-sharing networks and the cost and intensity of care in US hospitals. Med Care 50(2):152–160. https://doi.org/10.1097/MLR.0b013e31822dcef7

Bliss CA, Kloumann IM, Harris KD, Danforth CM, Dodds PS (2012) Twitter reciprocal reply networks exhibit assortativity with respect to happiness. J Comput Sci 3(5):388–397. https://doi.org/10.1016/J.JOCS.2012.05.001

Bollen J, Gonçalves B, Ruan G, Mao H (2011) Happiness is assortative in online social networks. Artif Life 17(3):237–251. https://doi.org/10.1162/artl_a_00034

Bonacich P (1972) Technique for analyzing overlapping memberships. Sociol Methodol 4:176–185

Cheadle JE, Stevens M, Williams DT, Goosby BJ (2013) The differential contributions of teen drinking homophily to new and existing friendships: an empirical assessment of assortative and proximity selection mechanisms. Soc Sci Res 42(5):1297–1310. https://doi.org/10.1016/J.SSRESEARCH.2013.05.001

Chen PY, Hero AO (2013) Node removal vulnerability of the largest component of a network. In: IEEE global conference on signal and information processing, GlobalSIP 2013—Proceedings, pp 587–590. https://doi.org/10.1109/GlobalSIP.2013.6736946

Corley HW, Chang H (1974) Finding the *n* most vital nodes in a flow network. Manag Sci 21(3):362–364. https://doi.org/10.1287/mnsc.21.3.362

Everett MG, Borgatti SP (2012) Categorical attribute based centrality: E-I and G–F centrality. Soc Netw 34(4):562–569. https://doi.org/10.1016/J.SOCNET.2012.06.002

Freeman LC (1977) A set of measures of centrality based on betweenness. Sociometry 40(1):35. https://doi.org/10.2307/3033543

Freeman LC (1978) Centrality in social networks conceptual clarification. Soc Netw 1(3):215–239. https://doi.org/10.1016/0378-8733(78)90021-7

Galinier P, Hamiez JP, Hao JK, Porumbel D (2013) Recent advances in graph vertex coloring. Intell Syst Ref Libr. https://doi.org/10.1007/978-3-642-30504-7_20

Ghalmane Z, Cherifi C, Cherifi H, El Hassouni M (2019a) Centrality in complex networks with overlapping community structure. Sci Rep. https://doi.org/10.1038/s41598-019-46507-y

Ghalmane Z, El Hassouni M, Cherifi H (2019b) Immunization of networks with non-overlapping community structure. Soc Netw Anal Min. https://doi.org/10.1007/s13278-019-0591-9

Grubesic TH, Murray AT (2006) Vital nodes, interconnected infrastructures, and the geographies of network survivability. Ann Assoc Am Geogr 96(1):64–83. https://doi.org/10.1111/j.1467-8306.2006.00499.x

Hollingsworth JM, Funk RJ, Garrison SA, Owen-Smith J, Kaufman SR, Landon BE, Birkmeyer JD (2015) Differences between physician social networks for cardiac surgery serving communities with high versus low proportions of black residents. Med Care 53(2):160–167. https://doi.org/10.1097/MLR.0000000000000291

Ibnoulouafi A, El Haziti M, Cherifi H (2018) M-centrality: identifying key nodes based on global position and local degree variation. J Stat Mech Theory Exp. https://doi.org/10.1088/1742-5468/aace08

Jia C, Li Y, Carson MB, Wang X, Jian Yu (2017) Node attribute-enhanced community detection in complex networks. Sci Rep 7(1):2626. https://doi.org/10.1038/s41598-017-02751-8

Krackhardt D, Stern RN (1988) Informal networks and organizational crises: an experimental simulation. Soc Psychol Q 51(2):123. https://doi.org/10.2307/2786835

Levit LA, Byatt L, Lyss AP, Paskett ED, Levit K, Kirkwood K, Schenkel C, Schilsky RL (2020) Closing the rural cancer care gap: three institutional approaches. JCO Oncol Pract 16(7):422–430. https://doi.org/10.1200/OP.20.00174

McPherson M, Smith-Lovin L, Cook JM (2001) Birds of a feather: homophily in social networks. Ann Rev Sociol 27(1):415–444. https://doi.org/10.1146/annurev.soc.27.1.415

Moen EL, Bynum JPW (2019) Evaluation of physician network-based measures of care coordination using medicare patient-reported experience measures. J Gen Intern Med. https://doi.org/10.1007/s11606-019-05313-y

Moen EL, Austin AM, Bynum JP, Skinner JS, O'Malley AJ (2016) An analysis of patient-sharing physician networks and implantable cardioverter defibrillator therapy. Health Serv Outcomes Res Method 16(3):132–153. https://doi.org/10.1007/s10742-016-0152-x

Moen EL, Bynum JP, Skinner JS, O'Malley AJ (2019) Physician network position and patient outcomes following implantable cardioverter defibrillator therapy. Health Serv Res. https://doi.org/10.1111/1475-6773.13151

Mollica KA, Gray B, Treviño LK (2003) Racial homophily and its persistence in newcomers' social networks. Organ Sci 14(2):123–136. https://doi.org/10.1287/orsc.14.2.123.14994

Newman MEJ, Clauset A (2016) Structure and inference in annotated networks. Nat Commun 7(1):11863. https://doi.org/10.1038/ncomms11863

Newman MEJ (2003) Mixing patterns in networks. Phys Rev E Stat Nonlinear Soft Matter Phys 67(2 Pt 2):026126. https://doi.org/10.1103/PhysRevE.67.026126

Pollack CE, Weissman GE, Lemke KW, Hussey PS, Weiner JP (2013) Patient sharing among physicians and costs of care: a network analytic approach to care coordination using claims data. J Gen Intern Med 28(3):459–465. https://doi.org/10.1007/s11606-012-2104-7

Rajeh S, Savonnet M, Leclercq E, Cherifi H (2020) Interplay between hierarchy and centrality in complex networks. IEEE Access. https://doi.org/10.1109/ACCESS.2020.3009525

Smith JA, McPherson M, Smith-Lovin L (2014) Social distance in the United States. Am Sociol Rev 79(3):432–456. https://doi.org/10.1177/0003122414531776

Tannenbaum SS, Soulos PR, Herrin J, Pollack CE, Xiao Xu, Christakis NA, Forman HP et al (2018) Surgeon peer network characteristics and adoption of new imaging techniques in breast cancer: a study of perioperative MRI. Cancer Med 7(12):5901–5909. https://doi.org/10.1002/cam4.1821

Tulu MM, Hou R, Younas T (2018) Identifying influential nodes based on community structure to speed up the dissemination of information in complex network. IEEE Access. https://doi.org/10.1109/ACCESS.2018.2794324

Valente TW, Coronges K, Lakon C, Costenbader E (2008) How correlated are network centrality measures? Connections (toronto, Ont.) 28(1):16–26

Zanghi H, Volant S, Ambroise C (2010) Clustering based on random graph model embedding vertex features. Pattern Recogn Lett 31(9):830–836. https://doi.org/10.1016/J.PATREC.2010.01.026

Zipkin R, Schaefer A, Chamberlin M, Onega T, O'Malley AJ, Moen EL (2021) Surgeon and medical oncologist peer network effects on the uptake of the 21-gene breast cancer recurrence score assay. Cancer Med. https://doi.org/10.1002/cam4.3720

Publisher's Note

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