Published in final edited form as:

Soc Sci Med. 2014 September ; 117: 67–75. doi:10.1016/j.socscimed.2014.07.014.

Results from using a new dyadic-dependence model to analyze sociocentric physician networks

Sudeshna Paul⁺, Nancy L. Keating[&], Bruce E. Landon[&], and James O'Malley^{*}

⁺Department of Biostatistics and Bioinformatics, Nell Hodgson Woodruff School of Nursing, Emory University, Atlanta, GA 30322, USA. sudeshna.paul@emory.edu [&]Department of Health Care Policy, Harvard Medical School, Boston, MA 02115-5899, USA

keating@hcp.med.harvard.edu; landon@hcp.med.harvard.edu ^{*}The Dartmouth Institute for Health Policy and Clinical Practice Geisel School of Medicine at Dartmouth Suite 300, Room 3066, 35 Centerra Parkway, Lebanon, NH 03766, USA James.OMalley@Dartmouth.edu Phone: 603-653-0854; Fax: 603-653-0896

Abstract

Professional physician networks can potentially influence clinical practices and quality of care. With the current focus on coordinated care, discerning influences of naturally occurring clusters and other forms of dependence among physicians' relationships based on their attributes and care patterns is an important area of research. In this paper, two directed physician networks: a physician influential conversation network (N=33) and a physician network obtained from patient visit data (N=135) are analyzed using a new model that accounts for effect modification of the within-dyad effect of reciprocity and inter-dyad effects involving three (or more) actors. The results from this model include more nuanced effects involving reciprocity and triadic dependence than under incumbent models and more flexible control for these effects in the extraction of other network phenomena, including the relationship between similarity of individuals' attributes (e.g., same-gender, same residency location) and tie-status. In both cases we find extensive evidence of clustering and triadic dependence that if not accounted for confounds the effect of reciprocity and attribute homophily. Findings from our analysis suggest alternative conclusions to those from incumbent models.

Keywords

Dyadic independence; Latent variables; Patient sharing; Physician influence; Sociocentric network; Transitivity

© 2014 Elsevier Ltd. All rights reserved.

Correspondence to: James O'Malley.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. INTRODUCTION

Professional physician social networks play a pivotal role in the diffusion of knowledge and adoption of new clinical practices (Burt, 1987; Coleman et al., 1966; Fennell & Warnecke, 1988; Nair et al., 2008; Valente, 1996). Understanding the topology and structure of physician professional networks in which ties between physicians depict potential information-sharing relationships can provide insights into how physician beliefs, behaviors and preferences are shaped. It may also reveal improved strategies for the spread of medical and healthcare information (West et al., 1999). To give a few examples: There is evidence suggesting that the structure of physician networks could be a factor underlying the cost and intensity of care in US hospitals (Barnett et al., 2012); Transfers from a peripheral to a more centralized hospital in a critical care network have been found to improve patient outcomes and survival through more efficient care (Iwashyna et al., 2009); Collaborations between physicians across specialties appears important to implementing therapeutic support protocols (Trzeciak et al., 2006); Physicians appear to rely mostly on their colleagues for information and advice about the care of their patients (Keating et al., 2007); Algorithms that detect naturally occurring “clusters” of physicians may be a basis for forming accountable care organizations (ACOs) (Landon et al., 2013).

In many of the above examples, physician networks have previously been studied using dyadic independent models that relate the probabilities of relationships existing between physicians (ties) to structural features of networks such as density, reciprocity (the phenomena whereby mutual relationships between physicians occur more commonly than by chance), and the distribution of the physician’s number of connections (degree) while accounting for heterogeneity between physicians (Keating et al., 2007; Landon et al., 2012). However, these and simpler models fail to account for higher-order effects (those involving 3 physicians) that account for clustering of ties within groups of three or more patients and phenomena such as transitivity (“a friend of a friend is a friend”). Statistical models that identify such effects have the potential to better inform research on the coordination of health care. In addition, determining whether the effects of reciprocity and homophily of physicians’ attributes are modified by phenomena such as transitivity may reveal mechanisms for achieving faster diffusion of information and improved quality and cost of healthcare.

In this paper we analyze two sociocentric networks of physicians’ professional relationships. The first network is a binary network of influential discussions on women’s health issues among physicians in a hospital-based practice. The second uses overlap of physicians patient cohorts to construct a network of professional relationships between physicians in a large geographical region. In both, the goal is to determine whether the effect of reciprocity differs between regions of the network with high and low prevalence of ties, and to estimate the independent effect of the similarity of physicians’ characteristics and of their attributes (whose effects constitute homophily) on their professional relationships.

The remainder of the paper is as follows. Below, we provide background material. In Section 2 we review and develop models while Section 3 describes estimation methods. In Section 4, we apply our model to a physician influential conversation network from a

Boston teaching hospital and a network of physicians formed on the basis of sufficient overlap of their patient cohorts from the North Dakota hospital referral region (HRR). An extensive discussion of results for both models follows. The paper concludes in Section 5.

1.1 Sociological Background

Sociocentric networks consist of collective information on the status of connections or “ties” between actors, often embedded within a complex social structure. The collection of physician “dyads”, “triads”, and other “clusters” in a physician network inform the social structure of a hospital or physician practice. Ignoring the hierarchy of social dependence (e.g. individual, dyadic, triadic) can lead to biased estimates of the effects of lower-level network features (e.g. within dyad features such as reciprocity). Recent analyses of physician networks have begun to focus on different aspects of the underlying social structure in addition to the actor-specific attributes as a means to understanding the behavior of the network as a whole (Lomi & Pallotti, 2012; Zappa, 2011). This has important implications for knowledge transfer and patient referrals; e.g., in a network with overlapping clusters or groups, information is more likely to diffuse faster than if the network contained non-overlapping groups. Similarly, the positioning of physicians can help explain patterns of ties and identify stochastic equivalence classes within the network, which in turn may provide valuable information about where best to seed a network-based intervention (e.g., coupons to purchase healthy living products) and the diffusion of information across the network (Hanneman & Riddle, 2005).

1.2 Network and Statistical Background

A common way of modeling sociocentric data is using p^* (exponential-family random graph) models (Frank & Strauss, 1986; Wasserman & Pattison, 1996). These treat the network as a single observation. An alternative model to the p^* model is developed from the premise of treating dyads as independent or conditionally independent random variables (Fienberg et al., 1985; Holland & Leinhardt, 1981; Yang & Yong, 1987). For example, the p_2 model for directed networks assumes dyadic independence conditional on actor specific random effects (van-Duijn et al., 2004). This approach is appealing as the dyad is the largest component of a sociocentric network such that knowing the status of one component does not restrict the possible states of any other component. Consequently, the conditional independence structure allows the model for the network to be induced from the model for a single dyad, ensuring desirable statistical properties (O'Malley, 2013).

Clustering induced by groups of 3 actors such as transitivity, stochastic equivalence, and structural balance are of great interest in characterizing the underlying social structure of the physician networks. Yet, it is only recently that conditionally dyadic independence models that include such terms have been developed (e.g., (Nowicki & Snijders, 2001); (Hoff, 2005; Hoff et al., 2002)). In related work, we combined the desirable features of the p_2 and p^* models to develop a longitudinal p_2 model that includes transitivity and other forms of between-dyad dependence as lagged terms (Paul & O'Malley, 2013). The latter model was notable because the dyad is the unit of analysis and reciprocity is the direct effect of the status of the tie $j \rightarrow i$ on the tie $i \rightarrow j$, arguably providing a more direct and intuitive measure of reciprocity. However, to date conditionally dyadic approaches are restrictive in that the

effects of reciprocity and triadic dependence were assumed to be additive (no interaction between them).

2. METHODS

In this section we adapt the longitudinal model we developed previously with the bilinear mixed effects (BME) model of Hoff (2005) to accommodate triadic dependence and other forms of between-dyad clustering in a cross-sectional network. Therefore, unlike the BME models developed by Hoff in which tie-status is the unit of analysis, under the resulting “extended p_2 model” (referred to as EP2 henceforth) developed here the dyad is the unit of analysis and reciprocity is the effect of an observed variable. This formulation allows more nuanced effects of reciprocity and more flexible adjustment for tie dependence in the network.

2.1 The p_2 model for sociocentric data

We first overview the p_2 model (see (van-Duijn et al., 2004) for a more detailed description than the following). Let $Y_{[n \times n]}$ be an adjacency matrix of connections (ties) between n actors in a directed binary network whose ij^{th} entry denotes the status of the relationship from i to j . The probability function representing the status of the dyad involving actors i and j , (Y_{ij}, Y_{ji}) , is given by:

$$\Pr((Y_{ij}, Y_{ji}) | \theta_{ij}, \rho_{ij}) = \text{gen.logit}^{-1}(\theta_{ij}y_{ij} + \theta_{ji}y_{ji} + \rho_{ij}y_{ij}y_{ji}), \quad (1)$$

where $\theta_{ij} = \mu_{ij} + \alpha_i + \beta_j$ and gen.logit denotes the generalized logit link function that transforms the dyad state-probabilities such that they can be related to a linear function of predictors. The parameter μ corresponds to density, ρ to reciprocity, and α and β to “gregariousness” (propensity of an individual to like others – in the sense of forming ties) and “popularity” (propensity of an individual to be liked by others) of the actors, respectively. The heterogeneity of the density and reciprocity parameters in (1) allows these effects to vary in magnitude across the network with actor-, dyad-, and tie-level covariates. The actor specific effects (α_i, β_i) are assumed to be random variables drawn from a bivariate normal distribution with mean $(0, 0)$, variances σ_α^2 and σ_β^2 and correlation $\rho_{\alpha\beta}$.

In (1) the status of Y_{ij} depends on $\rho_{ij}Y_{ji}$ on the logit-scale; the log-odds of $Y_{ij} = 1$ increases by ρ_{ij} if $Y_{ji} = 1$. The correlation $\rho_{\alpha\beta}$ is subtly different from reciprocity as it represents the extent that gregarious physicians are also attractive physicians. The assumption of conditional independence between dyads given the vectors of random effects (α, β) allows the model specified for a dyad to induce the model for the entire sociocentric network.

2.2 Bilinear mixed effects (BME) model

The BME model assumes conditional independence of ties Y_{ij} given tie-, dyad- and actor-level random effects (Hoff, 2005; Hoff et al., 2002). The probability of a tie is given by:

$$P(Y_{ij} | \theta_{ij}) = \text{logit}^{-1}(\theta_{ij}) = \gamma'x_{ij} + \alpha_i + \beta_j + \eta_{i,j} + \xi_{i,j} \quad (2)$$

where logit denotes the logistic link function, γ is the vector of covariate effects, and (α_i, β_i) are random effects (see (1)). Finally, $\eta_{i,j}$ is a tie-specific random effect with mean 0 and variance σ_η^2 . The correlation of $\eta_{i,j}$ and $\eta_{j,i}$ (denoted ρ_η) represents latent reciprocity; thus, ρ_η has a different interpretation to ρ_{ij} in (1).

Any unexplained variation in θ_{ij} in (2) not accounted for by observed covariates and the gregariousness and popularity random effects are modeled by a function, $\xi_{i,j} = f(v_i, v_j)$, of k -element latent vectors v_i and v_j (Hoff, 2002, 2005). The set $\{v_i\}_{i=1,n}$ can be interpreted as representing positions of actors in an unobserved “social” space with f a distance measure quantifying the “social closeness” of v_i and v_j . Some common examples of closeness measures for $f(v_i, v_j)$ are $-|v_i - v_j|$ (in the “distance” model), $\nu_i' \nu_j / |\nu_j|$ (in the “projection model”), and $\nu_i' \nu_j$ (in the “inner product” model). For a description of how these functions allow for transitivity and structural balance refer to Hoff (2005).

While continuous, binary or count observations are easily accommodated by generalizing (2) to include the appropriate link function and distribution for the data, thereby emulating generalized linear models, this is more challenging under (1) as the unit of analysis is bivariate. However, (e.g.) bivariate normal models for interval-valued ties or bivariate Poisson models for count-valued ties are natural adaptations.

2.3 Extended -model with latent effects (EP2 model)

We extend the traditional p_2 model in (1) to account for between-dyad dependence by incorporating the bilinear term from Section 2.2, using for illustration the inner-product form for f . We assume that the conditional probability for the status of the ij^{th} dyad satisfies:

$$\Pr((Y_{ij}, Y_{ji}) | \theta_{ij}, \rho_{ij}) = g^{-1}(\theta_{ij} y_{ij} + \theta_{ji} y_{ji} + \rho_{ij} y_{ij} y_{ji}), \theta_{ij} = \mu_{ij} + \alpha_i + \beta_j + f(v_i, v_j) \rho_{ij} = \rho + \tau f(v_i, v_j) \quad (3)$$

where $\{v_i : (v_{i1}, \dots, v_{ik})\}_{i=1,\dots,n}$ are independent random variables from a k -component multivariate normal distribution. Therefore, two sets of random effects – the 2-dimensional (α_i, β_i) representing “popularity” and “gregariousness” and the k -dimensional v representing an actor’s latent network (or “social”) position – account for between-dyad dependence. As in the p_2 model, $\rho = 0$ implies a purely additive model in which the presence of Y_{ij} does not directly depend on Y_{ji} . For undirected binary networks, $Y_{ij} = Y_{ji}$ and ρ_{ij} in (3) vanishes yielding the logit-link BME model in (2).

The EP2 model differs from the BME model as follows: (i) Reciprocity is modeled as an interaction effect between the constituent ties of a dyad as opposed to being a latent correlation; (ii) The bilinear term in EP2 is included in both the density and reciprocity components and so its effect is representing by two parameters (θ, ρ) in (3) as opposed to only impacting density. Within the EP2 density component, the bilinear term allows for clusters of higher and lower density of ties, as opposed to homogeneous density across the network. Within the reciprocity component, the bilinear term allows the network to contain clusters where reciprocated ties are relatively more and less common as opposed to constant reciprocation across the network. If τ is non-zero in then reciprocity is non-homogeneous.

The value of k (the dimension of the latent effects) is often determined as the smallest value for which there is no substantial lack of fit.

3. BAYESIAN MODELING AND ESTIMATION

Due to the complexity of the model, we use Bayesian methods for model fitting and estimation. First, to complete a Bayesian specification of the model, prior densities must be specified for the model parameters. The parameters of the prior distributions were chosen to ensure that information in the data was the primary determinant of the posterior distribution. Accordingly, multivariate normal prior densities with large variances were assumed for the latent effects (α , β , V) and the coefficients of the predictors in μ_{ij} or ρ_{ij} . Instead of specifying priors directly for $(\sigma_\alpha^2, \sigma_\beta^2, \rho_{\alpha\beta})$, the variances and correlation of the random effects (α_i, β_i) , we express the joint distribution of the random effects (α_i, β_i) in the product normal form (Cooper et al., 2007; Spiegelhalter, 1998). Finally, we assume that the actor's latent positions v_i have variance σ_v^2 and that σ_v^2 has an inverse gamma prior with mean = 1 and variance = 1. Details of the prior specifications and the Bayesian estimation procedure are provided in the Appendix.

4. ANALYSIS OF PHYSICIAN NETWORKS

We now introduce the professional physician networks: (1) a directed influence network of physicians in a primary care practice in a Boston area teaching hospital (Keating et al. 2007); (2) a directed physician network for the entire North Dakota hospital referral region (HRR).

4.1 Influence network of physicians in a primary care practice

Data—A physician network in a primary care practice at a major Boston teaching hospital ($N=33$) was measured as part of a study to examine how social networks influence information seeking behaviors regarding women's health issues (Keating et al., 2007). Different forms of this network have been analyzed by Keating et al. (2007) and O'Malley & Marsden (O'Malley & Marsden, 2008). In brief, the network was constructed from the responses of 33 out of 38 primary care physicians to a survey reporting the number of influential discussions (measured as 0, 1–3, or 4) about women's health issues they had with each other physician in the practice during the prior 6 months. We model these data as a directed binary-valued network where a tie ($Y_{ij} = 1$) indicates one or more discussions from actor i to j . Physician characteristics such as age, race, ethnicity, graduation year, number of years of practice in the Boston area, expertise within primary care, and percent female patients were also collected from the survey. Additional information such as sex, type of clinic, and number of ½ day clinic sessions per week were obtained from the hospital administrative files. Physicians were also asked to report “the one individual, inside or outside of the practice who is most likely to influence their practice regarding women's health issues”.

We include homophily measures of physician practice affiliation and gender in the model as dyadic covariates to account for the fact that physicians in the same practice and having the same gender may be more likely to form ties. Similarly, characteristics of the cited

physicians such as expertise in women health matters and total percent of women patients treated may also impact the extent that physicians initiate influential discussions. Thus, attributes of the cited physician's expertise in women's health, number of clinical sessions, and percent women patients were included as covariates in the density component of the model.

Descriptive network statistics—The average density of the physician influential conversation network is 0.154 with reciprocity approximately 0.19, implying mutual interactions between physicians were relatively more common than directional interactions. Figure 1 shows no significant association between the physicians' in-degree and out-degree distribution and as expected women health experts were consulted more than the other physicians (higher in-degree). Since the network is directed, there are 16 possible types of triads (Wasserman & Faust, 1994). Of these, we were especially interested in *three cycles* ($i \rightarrow h, h \rightarrow j, j \rightarrow i$), *transitive triads* ($i \rightarrow h, i \rightarrow j; h \rightarrow j$, or $j \rightarrow h$), *two-paths* ($i \rightarrow h, h \rightarrow j$) and *two stars* ($i \rightarrow h, i \rightarrow j$). To quantify the amount of global triadic clustering in the network we computed the proportions of cycles relative to two-paths, transitive triads to two-stars, and non-directional closed triads to non-directional two-paths. The resulting values 0.22, 0.36, and 0.29, respectively, suggest a high-degree of triadic clustering, justifying use of the EP2 model.

Results for key individual and contextual attributes—To evaluate the impact of the multiplicative bilinear term in the EP2 model, we compared the fit to that of the nested p_2 model. Table 1 displays the posterior mean and (2.5%, 97.5%) quantiles of the posterior distribution of the regression coefficients for the four covariates in the model. Results indicate that influential discussions were more likely to occur between physicians within the same clinic. For example, being in the same clinic compared to a different clinic increases the odds of tie formation by 5.9 (CI: 3.1, 11.8). Discussions were also more likely if the receiver physician is an expert in women's health and treats a higher percent of women patients. These findings are consistent with those from previous analysis of the data (Keating et al. 2007). The coefficient for gender homophily was significant, i.e., there was a higher likelihood of influential discussions between physicians of the same gender, an effect not previously identified by the p_2 model. We suspect that potential confounding with triadic clustering might have masked the true effect of gender homophily, rendering it insignificant when higher order dependence was ignored. The variance of the receiver random effect corresponding to physician "popularity" was bigger than that of their "gregariousness" possibly due to a few very influential women's health experts (e.g. physician 27 has in-degree 24). There is no evidence to conclude that popular physicians are also gregarious.

Results for the latent effects—We established that $k = 2$ was the optimal dimension of the latent position vectors on account that the fit did not substantially improve with higher dimensional latent vectors. The Bayesian posterior mean estimates of v_i for $i = 1, \dots, n$ from model without covariates are in Figure 2(a) while those adjusting for covariates, including the fact that within clinic ties are more likely, are in Figure 2(b). The gray lines indicate observed ties between actors. The latent positions estimated from the unadjusted model identify groups of physicians that primarily correspond to different clinics (e.g. the cluster

formed by actors 4, 10, 15, 23, 25, 30 corresponds to clinic 3). The covariate-adjusted model allows observed covariates of the actors (i.e., homophily measures of their similarity) to account for dependence in network ties before assessing if the remaining unexplained clustering is excessive (i.e., exceeds the amount expected by chance). The estimated variance, $\hat{\sigma}_v^2$ was smaller in the full model, implying a substantial reduction in social distances among physicians. However, substantial unexplained clustering remains.

An important finding from the original analysis (Keating et al 2007) was the extensive reciprocity (positive and large ρ). Specifically, if one physician cites an influential discussion with the second, then the second is also more likely to cite an influential conversation with the first (odds ratio 3.4). However, the additional of the bilinear term in the EP2 model makes the reciprocity effect small and insignificant, implying that reciprocity in this network may be an artifact of the presence of triadic and other higher-order configurations of actors. One theory is that individuals with reciprocated ties often have ties to other actors and thus triadic clustering manifests as reciprocity when between-dyad dependence is ignored. The point estimate of the slope of the bilinear term in the reciprocity component of the model, τ , is positive but not significantly different from zero implying there is insufficient evidence in support of a differential effect of the bilinear term on reciprocated than non-reciprocated ties; the evaluation of this test is novel feature of this paper made possible by the EP2 model. A reduction in the value of $\hat{\sigma}_v^2$ the estimated variances of the latent vectors, in the full EP2 model further justifies testing for this interaction effect. The EP2 model without the bilinear effect modifier in the reciprocity component found similar effects of the covariates but greater reciprocity compared to the full EP2 model, although the effect of the latter was still insignificant (Table 1).

Model evaluation—To evaluate the extent to which the model captures the presence of between dyad clustering of ties, we evaluated Bayesian predictive p-values for the proportions of: (i) cycles relative to two-paths, (ii) transitive triads relative to two-stars, and (iii) closed triads relative to connected triads. For a given statistic, the Bayesian predictive p-value is the posterior predictive probability that a value of the statistic at least as extreme as that observed obtains under the fitted model (Meng, 1994; Rubin, 1984). A value of 0.5 suggests that the model successfully captures the attribute quantified by the test statistic while values 0.1 or greater than 0.9 raise concerns about the appropriateness of the model (Gelman et al., 1996). The EP2 model performed better than the traditional p_2 model in capturing the transitivity in the network (Figure 3) with minimal evidence of model miss-fit in this regard.

Under the BME model with $k=2$ the latent positions of the actors were similar to those for the EP2 model (Figure 2). Furthermore, the estimated coefficients of the covariates from the BME model also show similar trends compared to the EP2 model. An estimated correlation of 0.5 between the dyadic latent effects suggests a general tendency towards reciprocation of ties – consistent with the findings for reciprocity under the EP2.

4.2 Physician professional network based on shared patients

Data—The second empirical analysis is of an excerpt of the physician networks identified from physician-patient encounter data of the 2006 Medicare beneficiaries in 51 hospital referral regions (HRR) across the US (Landon et al., 2012). The study sought to understand variation in network characteristics and patient sharing patterns among the constituent physicians based on physician, patient, and network level attributes, and identifying naturally occurring clusters of physicians based on shared care. The unipartite physician networks were discerned from a physician-patient bipartite system by connecting each pair of physicians who shared patients with one another.

In contrast to previous analyses of these data, we constructed a directed physician network by row standardizing the above sociomatrix, i.e., by (1) computing the relative number of shared patients between any pair with respect to the total number of patients shared by the originating physician, and (2) using a threshold value to binarize the proportions in (1). For consistency with prior work, the top 10, 15, 20 and 30th percentile values were used as potential thresholds. We report results based on the most conservative threshold that retains ties if the proportion was in the top 10% for that physician. This strategy is appropriate if a lot of ad hoc occur (e.g., a one-off appointment to a physician substituting for one's regular physician on a given day) that would create spurious ties under a larger threshold.

Physician age, gender, specialization, office location and patient age, sex, race, mean health status, and intensity of care were available. Physicians were classified as primary care physicians (PCPs) – including general internists, family practitioners, and general practitioners – medical specialists, or surgical specialists. The North Dakota HRR network, which consisted of 135 physicians, is analyzed here as we have a particular interest in comparing the results obtained under the EP2 model to prior results for this network as opposed to making comparisons among HRRs. The motivation for applying the EP2 model is in part due to the fact that the projection from the patient-physician bipartite network induces substantial triadic clustering, making results of prior analyses vulnerable to extensive bias arising from not accounting for such clustering. Similarity in physician attributes such as sex, specialization, and practice locations were included in the model as dyadic covariates; the coefficients of these terms represent specific forms of homophily.

Descriptive network statistics—The patient-sharing network had an overall density of 0.10 and a substantially higher rate of reciprocation, 0.21. While the observed sender degree distribution was symmetric, the receiver distribution was highly right skewed with a median of 6 and 95% of values within (0, 56). Exploratory analyses indicated that the observed proportion of cycles relative to two-paths, transitive triads to two-stars, and non-directional closed triads to non-directional two-paths were 0.22, 0.52, and 0.38, respectively. Therefore, a higher degree of triadic clustering appears to be present than would be expected by chance justifying use of the EP2 model.

Results for key individual and contextual attributes—When using the most conservative (top 10th percentile) thresholding scheme, similarity in practice specialty and gender was not significantly associated with physician ties (Table 2). Earlier results from

Landon et al (2012) based on all HRRs and using an undirected version of the network revealed that the likelihood of patient sharing is higher between males but lower between females compared to male-female dyads. We could have allowed separate baseline rates of tie formation for male-male, female-female and male-female dyads. However, because there were very few female-female (< 4%) dyads in the North Dakota HRR, we only included two dyadic variables for gender pairs (same gender and opposite gender). This might explain the lack of homophily on gender evident in the data. Physicians who practiced in the same locality were much more likely to have ties between them than physicians in different localities. The effects of the dyadic predictors (same sex, same practice location) on patient sharing relationship between physicians did not change across different thresholds (top 15, 20, 30, 40th percentile). However, the effect of “similarity in practice specialty” had a significant negative association under each of the higher thresholds, implying that physicians having complementary expertise are more likely to share patients compared to those having the same expertise.

Results for the latent effects—The latent factors reveal a large amount of structure in the network beyond what can be explained by a standard p_2 model (Figure 4a). There appears to be distinct separation of clusters of physicians in the peripheral region of the network while a large number of actors are lumped together in the central region forming multiple overlapping clusters and thus, less distinctive groups. Figure 4(a) depicts the actual observed ties between the actors overlaid on the social space. It is clear that the most central physicians belong to the same practice location (denoted by blue circles in Figure 4(a)). Accounting for the highly significant “practice location” covariate reduces the amount of clustering among estimated position of actors in this figure of the latent space although some localized clustering is still visible, indicating a potential interplay of both spatial and network proximity factors on physician ties. This is also evident from Figure (5(a–b)); although spatial proximity helped explain some of the clustering of physicians, the above covariates and other network features were not sufficient to account for all of the clustering in the network (Figure 5(a–b)). The slope parameter for the bilinear term, τ , was positive and marginally significant ($\tau \hat{0.24}$ (CI: 0.00, 0.52)), indicating that the closer two physicians are in terms of the latent variables (representing the social space), the greater the rate of reciprocity. However, because this effect is relative poorly identified, the difference in model fit between the full EP2 with the reciprocity – bilinear term interaction and reduced EP2 model without this interaction is not significant (Figure 4(a)).

Model evaluation—The parameter estimates from the P2, and the reduced and the full EP2 models are given in Table 2. We noticed a decaying effect in the reciprocity parameter (similar to what was observed in the physician influence network) when the bilinear term is introduced first in the density and then in the reciprocity component of the EP2 model (Table 2). Based on this trend, we believe that the reduction is due to unobserved factors accounting for triadic and higher-order forms of clustering among the ties also explaining some of the effect that otherwise manifests as reciprocity.

To evaluate the extent to which the EP2 model reproduces important structural features of the network, we evaluated Bayesian predictive p-values for the proportion of: (i) cycles

relative to two-paths, (ii) transitive triads relative to two-stars, and (iii) the in-degree distribution. A subtle but important feature of the data, the right skewed in-degree distribution, was well captured by the full EP2 model (Figure 5c). The EP2 model performed superior to the corresponding p_2 model in accounting for triadic clustering between physicians (Figure 5a–b). However, both models underestimated the high rate of transitivity in the data. Possible explanations are that the EP2 model with the 2-dimensional latent attributes or a symmetric bilinear term might be too restrictive to describe the high rate of triadic dependence in the network.

5. DISCUSSION

In this paper two important physician networks (a physician influence network and a physician patient care network) were analyzed using a new model to account for network effects such as transitivity, structural balance, and other forms of network clustering that have been ignored in prior analyses. Such prior studies on physician networks have mostly focused on the effects of individual factors, contextual characteristics and dyad-level effects on the presence of ties. Furthermore, there has been limited research on whether physician networks are shaped by the underlying topology: e.g., is there a tendency to form small, closed communities of peers, or for physicians to interact more globally with a few influential actors (e.g. opinion leaders) providing bridges between disconnected groups of physicians. Moreover, it has also remained unknown whether physician interactions can be explained by dyadic effects or if higher order complex network effects are required to account for the important features of physician networks.

Our key methodological contribution is the development of a model that allows for both the inclusion of a term to account for triadic dependence and also a term that allows for an interaction between reciprocity and triadic dependence. The resulting model thus allowed reciprocity effects to vary depending on the value of unobserved physician-level variables that account for triadic and other forms of between-dyad dependence by supposing the likelihood of a tie depends on the distance between the values of these unobserved variables in different physicians. The development of this model potentially solves the problems noted in the prior paragraph. Despite the addition of these complex terms to the base dyadic independence model (the p_2 model), the resulting model is still within the family of conditionally dyadic independent models ensuring convenient and well-behaved statistical properties.

The application of the new model to both physician networks studied in this paper yielded several interesting findings. While practice (clinic) and gender homophily were associated with a higher propensity of interactions among physicians in the physician influence network, expertise in woman's health and a higher percent of women among patients were important receiver physician characteristics associated with the initiation of influential discussions. These reinforce findings on tendencies of physicians to share information with and rely on peers that are proximal both organizationally and geographically (e.g. clinics, hospitals), and seek to consult with physicians who are deemed expert globally, leading to the emergence of potential "opinion leaders". While the former is potentially driven by opportunity and access, the latter points to an important role of leaders in adopting new

practices as well as transferring knowledge through the network and has huge implications in faster information diffusion (Clark et al., 2010; Valente & Davis, 1999).

One key finding was the diminishing effect of reciprocity when the bilinear terms were included as random effects first in the density component and then subsequently in the reciprocity component of the EP2 model (Table 2). This finding strengthens our belief that between dyad effects may be needed to account for dependencies in the formation of network ties, which if ignored may manifest as dyadic effects and lead to biased results. In the HRR physician network, proximity in practice location and complementary expertise were important factors that predicted the presence of ties among physicians. Therefore, the presence of physician ties in this network is shaped by physical proximity and the need for complementary as opposed to identical expertise in order to provide a general range of health care to patients. A marginally significant interaction between the bilinear term and reciprocity indicated that varying rates of reciprocity occurred among physicians based on their social proximity or presence in structurally equivalent parts of the network. To our knowledge, the ability to test this effect is unique to our EP2 model.

The patient-sharing physician network was more challenging to analyze due to the unusually high level of triadic clustering and skewed degree distribution. The higher propensity for triadic and higher-order dependence is an artifact of the network being a projection of a bipartite (two-mode physician-patient) network, which artificially inflates the number of closed triads (Wasserman & Faust, 1994). But this concern merely reinforces the need to use a model that accounts for between-dyad dependence in the presence of physician ties. By accounting for confounding from third-order effects, the empirical analyses of the physician networks presented in this paper are more trustworthy and informative than results obtained from earlier analyses of the same networks that ignore the extensive dependence between dyads.

Because the EP2 model is an extension of the generalized linear models family, it can be extended to fit networks with non-binary ties by using an alternative link function. For example, in the directed physician patient-sharing networks, instead of transforming the ties to binary values we could have used a Poisson model for the counts of shared patients and a log link to relate the expected number of ties to a linear function of the network features and covariates derived from observed characteristics of physicians. Another extension of our model would be to generalize the bilinear term to allow between-dyad clustering and stochastic equivalence, which are confounded in the model developed herein (Hoff, 2005), to be distinguished. A potential solution is to use an asymmetric (Hoff, 2009) or other more general representation of the bilinear term in the model. As in this paper, the representation of reciprocity as the direct effect of the reverse-direction tie on a tie's presence as opposed to a latent correlation and the interaction of this term with the bilinear mixed effect would make the resulting model distinct from existing bilinear models for sociocentric data.

In summary, we developed a novel method of modeling physician relationship patterns and used it to obtain new results for two important physician networks. We demonstrated the importance of adjusting for underlying structural dependence in order to estimate reciprocity and homophily effects. With the recent focus on coordinated health care, accounting for

dependence in network ties between physicians due to unobserved factors may yield more precise information on how best to improve knowledge transfer to aid faster adoption of good medical practices.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

REFERENCES

- Barnett ML, Christakis NA, O'Malley AJ, Onnela J-P, Keating NL, Landon BE. Physician patient-sharing networks and the cost and intensity of care in US hospitals. *Medical Care*. 2012; 50:152–160. [PubMed: 22249922]
- Burt RS. Social contagion and innovation: Cohesion versus structural equivalence. *American journal of Sociology*. 1987;1287–1335.
- Clark MA, Linkletter CD, Wen X, Miller EA, Mor V. Opinion networks among long-term care specialists. *Medical Care Research and Review*. 2010; 67:102S–125S. [PubMed: 20442338]
- Coleman, JS.; Katz, E.; Menzel, H. *Medical innovation: A diffusion study*. New York, NY: Bobbs-Merrill Company; 1966.
- Cooper NJ, Lambert PC, Abrams KR, Sutton AJ. Predicting costs over time using Bayesian Markov chain Monte Carlo methods: an application to early inflammatory polyarthritis. *Health economics*. 2007; 16:37–56. [PubMed: 16981192]
- Fennell, ML.; Warnecke, RB. *The Diffusion of Medical Innovations : An Applied Network Analysis*. Springer: 1988.
- Fienberg SE, Meyer MM, Wasserman S. Statistical analysis of multiple sociometric relations. *Journal of American Statistical Association*. 1985; 80:51–67.
- Frank O, Strauss D. Markov graphs. *Journal of the American Statistical Association*. 1986; 81:832–842.
- Gelman A, Meng XL, Stern H. Posterior Predictive Assessment of Model Fitness via Realized Discrepancies. *Statistica Sinica*. 1996; 6:733–807.
- Hanneman, R.; Riddle, M. *Introduction to social network methods*. Riverside, California: University of California, Riverside; 2005.
- Hoff PD. Bilinear Mixed Effects Models for Dyadic Data. *Journal of American Statistical Association*. 2005; 100:286–295.
- Hoff PD. Multiplicative Latent Factor Models for Description and Prediction of Social Networks. *Computational and Mathematical Organization Theory*. 2009; 15:261–272.
- Hoff PD, Raftery AE, Handcock MS. Latent Space Models for Social Networks Analysis. *Journal of American Statistical Association*. 2002; 97:1090–1098.
- Holland PW, Leinhardt S. An exponential family of probability distributions for directed graphs (with discussion). *Journal of the American Statistical Association*. 1981; 76:33–65.
- Iwashyna TJ, Christie JD, Moody J, Kahn JM, Asch DA. The structure of critical care transfer networks. *Medical Care*. 2009; 47:787–793. [PubMed: 19536030]
- Keating NL, Ayanian JZ, Cleary PD, Marsden PV. Factors Affecting Influential Discussions Among Physicians: A Social Network Analysis of a Primary Care Practice. *Journal of General Internal Medicine*. 2007; 22:794–798. [PubMed: 17404798]
- Landon BE, Keating NL, Barnett ML, Onnela J-P, Paul S, O'Malley AJ, et al. Variation in Patient-Sharing Networks of Physicians Across the United States. *The Journal of the American Medical Association*. 2012; 308:265–273.
- Landon BE, Onnela J-P, Keating NL, Barnett ML, Paul S, O'Malley AJ, et al. *Medical Care*. 2013; 51:715–721. [PubMed: 23807593]
- Lomi A, Pallotti F. Relational collaboration among spatial multipoint competitors. *Social Networks*. 2012; 34:101–111.

- Meng XL. Posterior Predictive P-Values. *The Annals of Statistics*. 1994; 22:1142–1160.
- Nair, HS.; Manchanda, P.; Bhatia, T. Asymmetric peer effects in physician prescription behavior: The role of opinion leaders. In: S. University. , editor. 2008.
- Nowicki K, Snijders TAB. Estimation and Prediction for Stochastic Blockstructures. *Journal of American Statistical Association*. 2001; 96:1077–1087.
- O'Malley AJ. The Analysis of Social Network Data: An Exciting Frontier for Statisticians. *Statistics in Medicine*. 2013 *In press*.
- O'Malley AJ, Marsden PV. Analysis of Social Networks. *Health Serv Outcomes Res Method*. 2008; 8:222–269.
- Paul S, O'Malley AJ. Hierarchical Longitudinal Models of Relationships in Social Networks. *Journal of the Royal Statistical Society: Series C*. 2013 *In press*.
- Rubin DB. Bayesian Justifiable and Relevant Frequency Calculations for the Applied Statistician. *The Annals of Statistics*. 1984; 12:1151–1172.
- Spiegelhalter DJ. Bayesian graphical modelling: a case-study in monitoring health outcomes. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*. 1998; 47:115–133.
- Trzeciak S, Dellinger RP, Abate NL, Cowan RM, Stauss M, Kilgannon JH, et al. Translating Research to Clinical Practice A 1-Year Experience With Implementing Early Goal-Directed Therapy for Septic Shock in the Emergency Department. *CHEST Journal*. 2006; 129:225–232.
- Valente TW. Social network thresholds in the diffusion of innovations. *Social Networks*. 1996; 18:69–89.
- Valente TW, Davis RL. Accelerating the Diffusion of Innovations Using Opinion Leaders. *The ANNALS of the American Academy of Political and Social Science*. 1999; 566:55–67.
- van-Duijn M, Snijders TAB, Zijlstra B. P2: A Random Effects Model with Covariates for Directed Graphs. *Statistica Neerlandica*. 2004; 58:234–254.
- Wasserman, SS.; Faust, K. *Social Network Analysis: Methods and Applications*. Cambridge University Press; 1994.
- Wasserman SS, Pattison PE. Logit models and logistic regression for social networks: I. An introduction to Markov Graphs and p^* . *Psychometrika*. 1996; 61:401–425.
- West E, Barron DN, Dowsett J, Newton JN. Hierarchies and cliques in the social networks of health care professionals: implications for the design of dissemination strategies. *Social science & medicine*. 1999; 48:633–646. [PubMed: 10080364]
- Yang YY, Yong GY. Stochastic blockmodels for directed graphs. *Journal of the American Statistical Association*. 1987; 82:8–19.
- Zappa P. The network structure of knowledge sharing among physicians. *Quality & Quantity*. 2011; 45:1109–1126.

Highlights

1. New methodology allows reciprocity to be modified by transitivity
2. Analyzed physician influential discussions and patient-sharing networks
3. Patterns of Influential discussions among physicians exhibit triadic dependence
4. Reciprocity among physician patient sharing networks modified by triadic effects
5. Geographically proximal physicians more likely to share patients

Figure 1.

Indegree versus outdegrees of physicians grouped by their expertise in women's health (squares denote women health, and circles other specialties) and clinic (black=1, red=2, green=3, blue=4) in the physician influential conversation network. Women health specialists were cited more frequently (e.g. physician 27 in clinic 4 is a WH-expert and had a very high indegree of 24) than others when discussing women health issues.

Figure 2.

Posterior mean estimate of physicians' latent positions in the social space. The latent positions are estimated from an EP2 model: (a) without covariates, and (b) after adjusting for the "clinic" covariate. They gray arrows denote the interactions between the physicians in the influential conversations network.

Figure 3.

Posterior predictive distributions of network characteristics from the physician influential conversations network (example 1) quantifying aspects of model-fit: the proportions of cycles among the two-paths, transitive triads among two stars and the total number of closed triplets (including cycles and transitive triads) among connected triplets (two stars or two paths) (L–R), respectively. The distributions are computed under the EP2 model (solid line) and the traditional model p_2 (dashed line). The solid dot represents the observed value of the statistic, revealing that the EP2 model much more successfully reproduces the triadic dependence in the network.

Figure 4.

Posterior mean estimate of the latent positions of the physicians in the “social space” of the patient-sharing generated network. (a) The latent positions are estimated from an EP2 model after adjusting for the all covariates. The different colors denote different location of practices and gray lines denote the observed patient-sharing relationship between physicians in the generated physician network. (b) comparison of posterior predictive distribution of the overall reciprocity estimate for the reduced EP2 and full EP2 model.

Figure 5.

(a–b) Posterior predictive distributions of overall triadic clustering in the physician patient-sharing generated network (example 2). The distributions are computed under our EP2 model (solid line) and the traditional p_2 model (dashed line). The solid dot represents the observed value of the statistic. (c) Posterior predictive distributions of physicians' in-degree distribution. The gray lines are degree distributions computed under the full EP2 model. The black line denotes the observed degree distribution of the nodes in the patient-sharing generated physician network.

Table 1

Parameter estimates for p_2 model and the reduced (no bilinear mixed-effects modifier of reciprocity) and full extended p_2 (EP2) model fit on the physician influential conversations network

Parameters	P2		EP2 (reduced)		EP2 (full)	
	Posterior mean	95% CI	Posterior Mean	95% CI	Posterior mean	95% CI
density (edges)	-4.181	(-4.868, -3.330)	-4.793	(-5.73, -3.76)	-4.673	(-5.620, -3.748)
<i>Nodal covariates</i>						
Cited physician Woman Health expert	2.151	(0.952, 3.273)	2.343	(1.113, 3.985)	1.781	(1.138, 2.469)
Cited physician %women patients	0.016	(0.001, 0.030)	0.017	(0.003, 0.032)	0.018	(0.004, 0.035)
<i>Dyadic covariates</i>						
Same gender	0.478	(-0.037, 0.954)	0.709	(0.188, 1.305)	0.629	(0.102, 1.097)
Same clinic	1.508	(1.075, 1.945)	1.775	(1.127, 2.464)	1.782	(1.138, 2.469)
reciprocity (mutuality)	1.223	(0.451, 2.131)	0.674	(-0.334, 1.555)	0.270	(-1.222, 1.440)
latent space modification of reciprocity	n/a	n/a	n/a	n/a	1.822	(-0.700, 5.884)
<i>Random effects Covariance parameters</i>						
sender variance	1.060	(0.511, 1.920)	1.1893	(0.544, 2.363)	1.171	(0.535, 2.306)
receiver variance	1.180	(0.616, 2.262)	1.438	(0.681, 3.012)	1.405	(0.647, 2.514)
correlation (sender, receiver)	0.028	(-0.461, 0.505)	0.077	(-0.521, 0.606)	0.102	(-0.415, 0.653)
Variance of latent attribute	n/a	n/a	0.747	(0.387, 1.231)	0.525	(0.202, 1.084)

Note: EP2 = Extended P2, and P2 = p_2 model.

Parameter estimates for P2, EP2 (no the bilinear mixed effects modifier term in reciprocity) and EP2 model fit on the shared-patient generated physician network based on a 10% threshold.

Table 2

Parameters	P2		EP2 (reduced)		EP2 (full)	
	Posterior mean	95% CI	Posterior mean	95% CI	Posterior mean	95% CI
density (edges)	-6.80	(-7.22, -6.42)	-7.40	(-8.18, -6.85)	-7.50	(-8.37, -6.99)
<i>Dyadic covariates</i>						
Same gender	0.06	(-0.11, 0.24)	0.06	(-0.11, 0.23)	0.10	(-0.01, 0.29)
Same specialty	0.05	(-0.07, 0.17)	-0.03	(-0.18, 0.11)	-0.04	(-0.19, 0.10)
Same practice locality	4.03	(3.77, 4.32)	3.28	(2.86, 3.80)	3.37	(2.99, 3.98)
reciprocity (mutuality)	3.20	(2.90, 3.45)	2.70	(2.36, 3.11)	2.43	(1.95, 2.93)
latent space modification of reciprocity	n/a	(n/a, n/a)	n/a	(n/a, n/a)	0.24	(0.00, 0.52)
<i>Random effects Covariance parameters</i>						
sender variance	2.24	(1.69, 2.86)	1.37	(0.91, 1.97)	1.47	(1.00, 2.33)
receiver variance	4.10	(3.11, 5.51)	6.61	(4.78, 8.99)	6.91	(4.89, 9.83)
correlation (sender, receiver)	-0.51	(-0.63, -0.34)	-0.76	(-0.88, -0.58)	-0.71	(-0.84, -0.51)
variance of latent effects	n/a	(n/a, n/a)	1.29	(1.07, 1.55)	1.20	(1.00, 1.45)

Note: EP2 = Extended P2, and P2 = p_2 model.