



## Information Systems Research

Publication details, including instructions for authors and subscription information:  
<http://pubsonline.informs.org>

### Network Dynamics: How Can We Find Patients Like Us?

Lu (Lucy) Yan, Jianping Peng, Yong Tan

To cite this article:

Lu (Lucy) Yan, Jianping Peng, Yong Tan (2015) Network Dynamics: How Can We Find Patients Like Us?. Information Systems Research 26(3):496-512. <https://doi.org/10.1287/isre.2015.0585>

Full terms and conditions of use: <http://pubsonline.informs.org/page/terms-and-conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact [permissions@informs.org](mailto:permissions@informs.org).

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2015, INFORMS

Please scroll down for article—it is on subsequent pages



INFORMS is the largest professional society in the world for professionals in the fields of operations research, management science, and analytics.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

# Network Dynamics: How Can We Find Patients Like Us?

Lu (Lucy) Yan

Department of Operations and Decision Technologies, Kelley School of Business, Indiana University,  
Bloomington, Indiana 47405, [yanlucy@indiana.edu](mailto:yanlucy@indiana.edu)

Jianping Peng

Sun Yat-sen Business School, Sun Yat-sen University, Guangzhou 510275, China, [mnsjp@mail.sysu.edu.cn](mailto:mnsjp@mail.sysu.edu.cn)

Yong Tan

Michael G. Foster School of Business, University of Washington, Seattle, Washington 98195, [ytan@uw.edu](mailto:ytan@uw.edu)

Social networks have been shown to affect health. Because online social networking makes it easier for individuals to interact with experientially similar others in regard to health issues and to exchange social support, there has been an increasing effort to understand how networks function. Nevertheless, little attention has been paid to how these networks are formed. In this paper, we examine the driving forces behind patients' social network formation and evolution. We argue that patients' health-related traits influence their social connections and that the patients' network layout is shaped by their cognitive capabilities and their network embeddedness. By studying longitudinal data from 1,322 individuals and their communication ties in an online healthcare social network, we find that firsthand disease experience, which provides knowledge of the disease, increases the probability that patients will find experientially similar others and establish communication ties. Patients' cognitive abilities, including the information load that they can process and the range of social ties that they can manage, however, limit their network growth. In addition, we find that patients' efforts to reach out for additional social resources are associated with their embeddedness in the network and the cost of maintaining connections. Practical implications of our findings are discussed.

**Keywords:** social networks; healthcare; network dynamics; homophily; cognitive capabilities

**History:** Michael Smith, Senior Editor; Karthik Kannan, Associate Editor. This paper was received on April 23, 2011, and was with the authors 29 months for 4 revisions. Published online in *Articles in Advance* August 11, 2015.

## 1. Introduction

Social networks have been found to affect health. Research on social networks' effects on health, which emerged as early as the 1970s, demonstrated that social networks have the potential to affect mortality (Cassel 1976, Cobb 1976, Berkman and Syme 1979). More recent studies have documented that patients' social networks influence their perceptions of healthcare experience and conceptions of their roles in their own care, reshape their healthcare behaviors and coping strategies with illness, and affect their navigation of search services (Smith and Christakis 2008; Borgatti et al. 2009; Thoits 2010, 2011; Gage 2013). Research on online social networking has found that the ease of exchanging social support—assistance and protection given to others (Cobb 1976, Shumaker and Brownell 1984)—enables people to shift from the role of passive recipient to that of active consumer of health information (O'Grady et al. 2008, Fichman et al. 2011). Participation in online social networking activities helps individuals gain better disease management

knowledge and achieve better health conditions (Frost and Massagli 2008, Yan and Tan 2014).

The utilization of social media in healthcare is growing (Agarwal et al. 2010). Among social media applications, online healthcare communities, where patients are information contributors as well as consumers, are particularly active (O'Grady et al. 2008, Fichman et al. 2011). These online health networking sites help individuals take active control of their health as well as seek and receive help from sources such as physicians, nurses, other professional healthcare providers, and peers, i.e., other patients with similar illnesses. Patients can choose what to share and with whom to interact as they build their own way to contribute to and to draw on the value of the "wisdom of patients." Social media-enabled platforms make it easy for patients to write what they want and broadcast it instantly to their friends and followers. Simultaneously, others can "like," "follow," "link," and "tag" content shared by other patients or users (PwC HRI Social Media Consumer Survey 2012).

The benefits are obvious. Hundreds of thousands of individuals exchange advice and experiences in these virtual communities (Preece and Ghazati 1998, Wang et al. 2008). Members of these healthcare-focused virtual spaces can raise questions, search disease information, share medical records, discuss problems, get second opinions from others who suffer from the same disease, and compare self-outcomes with a larger population. As more websites bring the social networking revolution to healthcare, online healthcare communities have gained significant social, organizational, and economic importance.

Despite the obvious benefits of online healthcare communities, massive amounts of information could cause overload problems for information seekers. This matter can be compounded when information seekers cannot find information relevant to their needs. Therefore, there is great value in the ability of patients to build their own network with other patients, especially with ones who have similar medical conditions, so as to more easily find relevant information, exchange social support, and receive expected social gains from their online social interactions and other activities.

Although network phenomena have become more prominent in health research, network dynamics over time is an often-overlooked area of network research (Borgatti et al. 2009). Most conceptual and empirical attention has concentrated on the effects of social support. However, comprehending social networks' effects on health requires not only an understanding of how networks are functioning but also how they are formed (McPherson et al. 2006, Smith and Christakis 2008). Most importantly, as the vehicle through which social support occurs, social networks provide the environment for social assistance and protection that may not otherwise be available. Despite calls for research on social networks (Berkman and Glass 2000), possibly owing to the difficulty of seeing whole groups of individuals and their interconnections at once, there is limited research on their formation and evolution.

In this paper, we are therefore interested in the networks themselves, and we focus on the social relations surrounding an individual. By contrast to social support studies, which assess the effects of the quality or quantity of an individual's already existing social ties, we explicitly examine the formation of these ties. In particular, we look at a healthcare community that focuses on mental problems, study how patients identify other experientially similar patients, and select the proper ones to establish social connections. We focus on two plausible mechanisms in the formation of a social tie: homophily and cognitive capabilities. Research has shown that strategic selection behavior

is a result of homophily (i.e., a preference for similarity), a concept that social scientists have identified as one of the important mechanisms for why people create, maintain, dissolve, and reconstitute communication networks (Monge and Contractor 2003). To reexamine the effects of homophily in the context of this study, we constructed a set of multilevel nested measures that concern similarities in disease, symptoms, drugs, and treatments. We find that the similarity in firsthand disease experience with respect to prescription drugs increases the probability that patients will find experientially similar others and establish communication ties.

We also find evidence that shows that individuals' actions in the online healthcare community are constrained by their cognitive abilities. Whereas patients' communication efforts are restrained by information quality and quantity, their sustainable interactive expansion in the virtual environment is also constrained by the social network in which they are embedded, as complex network structures shape the spread of information, norms, and social exchanges. These information exchange activities can affect subsequent medical decisions, which, in turn, enhance or weaken patients' gains from online health communities and eventually determine whether they continue to participate, as well as the success and effectiveness of social networking in the healthcare context. To characterize a patient's cognitive capacities, we constructed various measures, including network size, network cost, and network search cost. By explicitly accounting for individuals' characteristics and their network positions, as well as the unique features that pertain to the healthcare domain, such as illness and medication, we add to the literature by presenting evidence of how individual constraints, resulting from network endogenous features, affect social network evolution in the context of the online healthcare community.

Our work also has practical implications. Although social media has been used in healthcare in a variety of ways, researchers believe that it could have an even stronger role in treating socially shaped diseases, including obesity and depression. In social disease, a recent concept, a substantial proportion of the burden of disease has the potential to be directly mediated by social networks (Coiera 2013). Evidence suggests that, because individuals' behaviors can be reshaped by the behaviors of social contacts and social ties, health-related emotional status, such as depression or suicidality, can spread through networks (Smith and Christakis 2008). Our work, by examining the social relations of patients with mental problems, thus falls into the area of social disease studies and provides an understanding of how these social ties and consequent networks are formed. As a result, our

findings have implications for the possible development of network therapy and other effective network interventions.

Section 2 presents our theoretical arguments and research hypotheses. We then describe our research method in §3. The data and variables used in our model are explained in §4. We present the results and findings in §5. Finally, in §6, we conclude the study and discuss practical implications.

## 2. Conceptual Background and Research Hypotheses

We examine an online healthcare community that offers services ranging from a basic tier of emotional support and information sharing to quantified self-monitoring for patients with mental problems. The community also provides technology tools, such as collaborative filtering, to help patients more easily identify others who are experiencing the same illness (experientially similar others) to share their health information, discuss their problems, and obtain advice.

Patients are drawn to an online healthcare network because it has value for them (Kadushin 2012). Individuals' access to resources and valued social attributes embedded within their social networks indicates that individual-level social capital overlaps with social support and social networks (Scheffler et al. 2008). Social capital is defined as "the resources embedded in social networks that are accessed and used by actors for actions" (Lin 2001, p. 25). A similar perspective is revealed in Nahapiet and Ghoshal (1998), who explicitly emphasized the difference between social capital and other forms of capital, noting that the resources of social capital are embedded in a social structure and in the relationships between individuals and their connections within communities (Putnam 1995). In this online healthcare community, patients share their health information and support each other. These collaborative behaviors thus constitute the foundation for two types of resources: personal resources, provided by shared health experiences, and social resources, provided by online connections (Lin 2001). In this section, we begin our exploration of the motivational foundations for patients' social networking behavior and then examine the cognitive aspect of social networking.

### 2.1. Social Network: Motivation and Dynamics

Typically, there are two complementary aspects of social networks in community and communication structures: the connections between some individuals and the nonconnections between other individuals. These correspond to two kinds of human motivations in regard to the social environment. One motivation

is to stay within one's social cocoon for one's connections. The focal person is typically called an "ego," and her social connections are called "alters." When a patient (the ego) needs help, her first behavior would be to check her existing social connections (alters) for immediate support. As explained by relational psychoanalytical theory (Greenberg 1991), the patient feels safety, comfort, and support from her current social relations. As a result of such support, she feels even closer to her existing connections, which contributes to a dense egocentric network.

The other motivation is to reach out and make connections where there were none (Kadushin 2012). If information or support is not immediately available in a patient's existing social relations, then the patient has a need to reach out to other patients and to establish connections with them (connect to more alters) so that she can acquire new social resources. If a patient has connections to many other patients, she is able to access more resources and, thus, becomes more efficacious. These two motivations are fundamental forces that drive the dynamics of individuals' network structure. Specifically, in the process of meeting patients' various social needs, regardless of the type of support required, the network continues to develop as social ties are established or idled; this raises the issue of the role of selection and homophily.

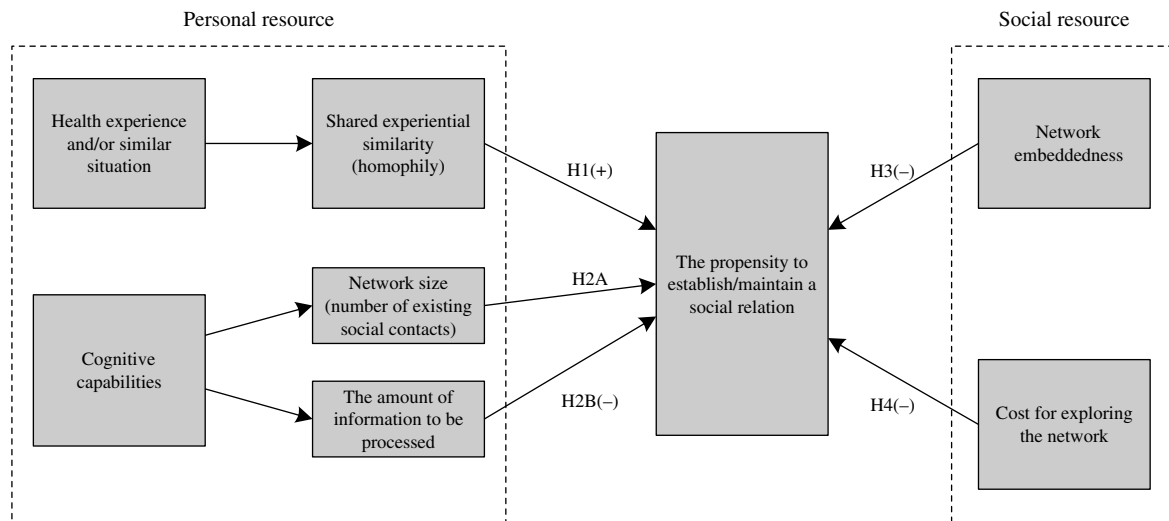
Social networks are not static but, rather, evolve over time (Snijders 2001, 2005). A social network is the structure of an interactive process, whereas social support is the function (Hinson Langford et al. 1997). By establishing new connections and removing idle ones, patients adjust their information management as well as develop their ability to identify and process useful information. This learning process is demonstrated through dynamic changes in the network. As stated by Snijders (2005), these changes could be caused by the mechanisms related to individuals' characteristics (network exogenous characteristics) and by purely structural network endogenous mechanisms. In addition, a patient's attempt to meet her social needs causes a trade-off between comfort and support derived from her egocentric network and the benefits achieved by forging bridges to a wider network. The dynamics in social relations reflect the extent to which social support can be derived from the network individuals. In the remainder of this section, we will discuss the different mechanisms used to develop social relations in meeting patients' social needs. Figure 1 presents the conceptual framework with the two dimensions of social capital used in our study and the associated hypotheses.

### 2.2. Selection and Homophily

The best-known driving force for forming or removing a relation is homophily, a widely used term that



Figure 1 Conceptual Framework



refers to the tendency of individuals to seek out similar others. Researchers have found that sharing an attribute produces interpersonal attraction (Reagans 2005) and affects the strength of communication. Proposed by Byrne (1961, 1971), the similarity attraction hypothesis indicates that the more similar people are, the more they are attracted to each other. Such attraction has powerful implications for people's social worlds, including the type of information they receive and the interactions they experience (McPherson et al. 2001).

Connections in online healthcare communities do not necessarily have geographic proximity, and, thus, the ability to remotely access resources facilitates communications. In addition, the choice to be anonymous offers patients the opportunity to ask for information that they might otherwise find difficult to ask about. More importantly, experientially similar others with shared experiences, particularly in an online healthcare community formed by patients with comparable health conditions, are thought to offer specialized support tailored to be more helpful than the support available from significant others (Gage 2013). By contrast to individuals who are similar based on common social characteristics, such as age, gender, or demographic features, patients with firsthand experience or knowledge in coping with a particular health issue can validate emotions and concerns and offer meaningful advice (Thoits 2011). This results in the selection of characteristics or attributes in the given situation for homophily rather than for widely studied ascribable features, such as age or gender, or acquired features, such as marital status or occupation. In regard to selection, health-related traits, or even health status, can contribute to the creation or dissolution of specific network ties or to the formation of a network (Smith and Christakis 2008). For example, Preece

and Ghazati (1998, 2001) examined the contribution of empathy, also known as emotional support, to the online community and found that people with similar backgrounds, in terms of illness, addiction, disability, and other comparable health experiences, were more likely to form relationships.

Online health communities are not a substitute for local healthcare services but rather a supplementary outlet when patients lack the social support they need. Patients turn to online healthcare communities to find support and exchange health information, which has been shown to be the most sought after form of social support in an online healthcare community (Yan and Tan 2014). This suggests that health information and, in particular, disease experience, is what healthcare community members care most deeply about. Further, it is the desire to learn about these personal experiences (e.g., similar symptoms suffered, medication problems faced) that triggers communication among patients (Wang et al. 2008). Different from the observational learning seen in other social contexts, *experiential learning* is based on experience (O'Grady et al. 2008). This type of informal health education, such as the revelation of personal anecdotes or off-label usage experiences as well as the access to "patients like me," increases patients' capacity for collaborative education.

The network formed by experientially similar others also provides the opportunity for social comparison (Thoits 2011). According to social comparison theory (Swann and Brown 1990), individuals develop their self-concept by comparing themselves to others in their social network. Patients may observe other patients who have faced similar health issues and compare their self-reaction and coping styles with those whom they believe to be coping well or, by contrast, compare themselves against less fortunate

similar others as a means to avoid potential problems. Information, resources, and strategies for getting the most out of their healthcare can be exchanged through their social relations. Such a social comparison process, therefore, plays an important role in the development of a patient's social network by allowing the patient to select the proper similar other patients from whom to receive as well as to give social support to as a means to enhance their coping abilities, emotional adjustment, and self-esteem. The patient's engagement in the process of social comparison and exchange of social support is also indicative of the patient's social competence, which plays an essential role in the formation of social relationships (Stewart 1993). Based on these concepts, we believe that patients' firsthand experiences with the same health conditions contribute to the concept of homophily, and we propose our first hypothesis:

**HYPOTHESIS 1 (H1).** *The more experientially similar a patient (alter) is compared to the focal patient (ego), the more likely that the ego will form a social connection.*

### 2.3. Cognitive Capabilities

A patient's acquired larger network continues to develop with the aid of additional functionality tools provided by the healthcare community. As a critical mechanism for facilitating patients' information seeking in online healthcare models, search or collaborative filtering enables the identification of potentially related patients in terms of similar situations or experiences (Eysenbach 2008, Swan 2009). It is becoming increasingly popular for a patient, based on other patients' experience, to search for potentially related issues that she might encounter. An example of such an issue is the various side effects reported by a large pool of patients who are undergoing a certain treatment. In addition, useful information or support may not always be embedded in the most recent message or contact. Thus, patients not only process new discussions but also browse other patients' health histories to assess their similarity or relevance.

Despite the obvious benefits of communicating with more patients who suffer from similar health conditions, social interactions involve the strain of dealing with other individuals as well as the time and effort to contribute their own experiences and draw from the collective knowledge (Jones et al. 2004, Wasko and Faraj 2005). In addition, there are cognitive limits, such as the size of the network and the amount of information, to what individuals can manage and process effectively. Moreover, patients' social embeddedness in the network, which is related to the cost of their social activities and community engagement, also affects the type of social support that can be obtained and the extent to which the social support is available. As a result, the cognitive effort needed

to search and compare useful resources reshapes the structure of a patient's social network.

**2.3.1. Cognitive Limits and Information Processing.** When patients join an online healthcare community to gain useful information and to communicate, they have the opportunity to meet with many others who are like-minded. Having an immediate connection to many other patients allows a patient to obtain support in a timely manner and receive advice from various sources who hold potentially different opinions. Thus, expanding the network size, that is, the number of direct social contacts, is a good way for patients to integrate themselves into the online environment. However, as their network grows larger, patients will encounter the problem of "too many contributors," which is increasingly common in newer social media platforms (Kane 2011) and challenges online users to find a satisfactory approach to information acquisition and use (Bawden and Robinson 2009). This problem is a consequence of cognitive psychological processes that limit the extent of human actions in terms of reaching out to other parts of the network (Kadushin 2012). Classic studies have found that the average person has a maximum effective network size of around 150 (Zetterberg 2011) and that the number of close relationships that individuals can manage is around 10 to 20 (Parks 2007). Although a recent study (Kadushin 2012) has found, in certain situations, an increased number of connections that an individual person can effectively manage, the limitations imposed by cognitive abilities still held.

The above arguments suggest a curvilinear relationship between patients' network expansion decisions and their network size. When a patient's egocentric network is small, there is a higher probability for her to establish more social ties to access more social resources. However, once the network size reaches the patient's cognitive limit, the probability of establishing new connections will decrease. Thus, we expect an inverted-U relationship between a patient's network size and the probability of establishing new social connections.

**HYPOTHESIS 2A (H2A).** *A patient's network expansion has a curvilinear (inverted-U) relationship with network size.*

Apart from the cognitive limits on managing network contacts, social media and the exponential growth of easy access to health information provide seemingly endless opportunities to inform and teach patients (Cline and Haynes 2001). Whereas much research has focused on the quality of information, other research has shown that the quantity of health information can be a barrier for patients. Too much information, along with too much diversity of the information, can lead to overload, especially as

web-based information sharing becomes more central to everyday life (Bawden and Robinson 2009).

This “overload” phenomenon is not new. First mentioned by Gross (1964), the problem of “information overload” is widely recognized today in various contexts (Edmunds and Morris 2000). In general, information overload occurs when the amount of information received becomes a hindrance, rather than a help, even though the information is potentially useful (Bawden and Robinson 2009). In particular, information overload may make it difficult to accurately identify relevant cues, resulting in decreased performance. The negative consequences of information overload include illness, stress, loss of job satisfaction, and ineffective actions (Edmunds and Morris 2000, Denning 2002); difficulty in obtaining useful and relevant information, even though there is an abundance of information available (Yellowlees and Brooks 1999, Edmunds and Morris 2000); a reduction in the likelihood that individuals will read long enough to discover potentially valuable contributions (Weiss et al. 2008, Jones et al. 2004); and patients’ being overwhelmed by too many choices (Bawden and Robinson 2009).

Even though the features of collaborative filtering provided by the online healthcare community make it easier for patients to find experientially similar others, the massive volume of information and the assessment of its quality are still prominent issues (Christensen et al. 2004). In addition, the most up-to-date information is not necessarily what patients need. Sometimes, patients need to understand their health conditions that reside in older messages, thus making new content-oriented collecting technologies (e.g., RSS) less effective, as they constantly refresh with the newest information and purge the old content. Another common problem in regard to health information is that patients may not know exactly what they want or which keywords to use, even if the information is available. With so much information available and so many similar patients with whom to communicate, patients may experience anxiety about whether they have missed an important piece of information. This is of particular concern to patients with a life-threatening illness, because obtaining a critical piece of information could be a matter of life and death. Therefore, based on the understanding that a patient’s network is constrained by her cognitive ability, we expect that it is less likely for a patient to go beyond her current social relations for social support when she has more health conditions to take into consideration.

**HYPOTHESIS 2B (H2B).** *The more experiential information that a patient needs to process, the less likely that she will establish new social connections.*

**2.3.2. Network Embeddedness and Opportunity Cost.** According to Kadushin (2012), social support includes the structure of the network within which an individual is embedded. As an antecedent of social support, some degree of connectivity must exist to derive the desired social support from the environment (Hinson Langford et al. 1997). Network embeddedness (Barrera 1986) is the connectedness of a patient to other members of an online healthcare community (Grewal et al. 2006, Ransbotham et al. 2012) and is often referred to as her “position” in a social network (Lin 2001), which determines the extent and diversity of social support that she can access. In online healthcare communities, the exchange of social support is initiated by patients’ sharing their health condition and experience. One of the values of sharing this information lies in the opportunities it provides to discuss it with other patients. For someone who was recently diagnosed with an illness or who has started a new treatment, learning from patients who have been through the disease and can discuss treatment effects and prognosis will help her to better understand what to expect and to be prepared.

When individuals communicate and collaborate, information and knowledge can be transformed (O’Grady et al. 2008). The wisdom of patients emerges as individuals connect on the basis of similar health characteristics and then learn from experiential stories about how their disease affects their lives. When this diversified “wisdom” is combined with existing information, the result is a richer profile of a disease and its manifestations as well as the generation of “new” knowledge. Patients in a more central position are more likely to be in the middle of, and hence benefit more from, the exchanges in the social network. In other words, the more embedded a patient is, the greater access that she already has to the health knowledge and experience that has been exchanged among other patients (Ransbotham et al. 2012) and, hence, the diminishing benefits of reaching out for new social contacts.

**HYPOTHESIS 3 (H3).** *The more embedded a patient is in the network, the less likely that she will establish new social relations.*

Aside from the cost of Internet access, the cost of accessing information in online healthcare communities includes the cost of time spent searching for information and the risk of obtaining inaccurate information (Bundorf et al. 2006). Although the monetary cost of accessing health information in these online healthcare communities is low, the opportunity cost of searching for resources as well as of obtaining benefit from individuals’ online actions is likely to be uncertain, especially when a patient is looking for social support that is not available within her immediate



contacts. The reasons are twofold. First, because of the option to be anonymous, there is no effective quality control of user-generated content in virtual spaces, and patients might obtain inaccurate or even potentially dangerous information (Christensen et al. 2004). In the area of health literacy, research on the collaborative environment has raised concern that these online places could be unfriendly when individuals express dissenting ideas or views (Gruhl et al. 2006). Therefore, identifying a random patient by using search tools can have an uncertain cost for a patient when verifying related information and assessing whether there is a “mismatch.”

Second, other research has addressed the issue of potential harm from misinformation and has suggested that an online healthcare community is a “buyer beware” environment (O’Grady et al. 2008). The expected benefit of seeking health information is based on the likelihood that the information has value and will improve the individual’s health. Under such circumstances, patients in the online healthcare community need to have a sense of comfort in regard to their actions and social relations. This sense of comfort includes not only supportive relations but also the reliable social resource embedded in the connection. However, it takes time to develop the sense of reliability between a pair of unknown patients. Patients can, instead, use the structural component of a network to facilitate linkages between people and reduce the potential costs (Uphoff 2000). For instance, friends of friends can be used as a third-party means of reliable social connections. Nevertheless, there is also a transaction cost associated with this approach. When the cost is low, it is more likely that a patient will find other patients and possibly construct connections; however, if the cost is high, the probability of identifying similar others would decrease and, consequently, lower the chances of establishing social connections. Therefore, we expect a negative relationship between the cost incurred in a patient’s network expansion and the probability of establishing new social connections.

**HYPOTHESIS 4 (H4).** *The probability of establishing new social relations decreases with a patient’s cost to explore the network.*

### 3. Research Setting and Method

#### 3.1. Research Setting

Health-related social networking sites vary from the disease-oriented (e.g., breast cancer) to multitasking portal sites that contain information that pertains to different sectors of the health field (e.g., health professionals, medications, patients). As noted, we are interested in social media applications among patients,

and this study focuses on a patient-centric online environment. Thus, we collected data from a Health 2.0 website that enables patients to construct patient-patient dyads to improve their health education. This disease-based healthcare community provides a social platform for patients to share personal health information, exchange medical experiences and knowledge, and learn from one another to improve their own health management. There are two main channels for individuals to use in this online healthcare community. The first channel consists of patients’ profile pages. At the time of registration with the website, patients need to disclose their disease conditions and create their own health profiles. Based on the category of the disease, patients are assigned to the corresponding disease-based community, where they can meet other patients with similar health problems. Personal Web pages contain a health profile for each member, including personal and demographic information, health condition history, outcomes, symptoms, counseling or therapy, and detailed records of treatments. Patients can update their current health condition at any time by answering a few questions. A chart provides a direct interpretation of each patient’s current condition and contains links to her previous input to show her health condition history and changes. Such profiles are publicly available and shared with focal members within the community but are only partially observable to those outside the specific disease community.

Patients also can find and communicate with similar members in the forum, where they discuss various topics, share concerns about their disease, and describe treatments or experiences. This channel possesses functions similar to those in chat rooms but without limiting access to community members. The website provides a subscription tool that helps patients to exchange information with members whom they select and keep track of their disease progress. The tool allows patients to receive updates about profile information, treatment changes, or related topics. From the subscriber and subscription lists, a patient can thus construct her network and gradually embed themselves in the community. Patients are therefore networked in the community through these connections.

#### 3.2. Research Model

Social network data involve complex structures that include a set of people, their characteristics, and their pairwise relationships (Goodreau et al. 2009). The structure of such social relations is defined by both the distribution of individual characteristics and the dynamics of interactions. The exponential random graph (ERG) model that portrays the characteristics of the focal network structure by comparing it with all



alternative layouts was perfectly suited to our study. That is, this method specifies and compares the probability of the current network configuration with all alternative structures for the given set of ties and a set of individuals and their attributes. It estimates the coefficients and determines the impact of statistical features for a specific set of data. However, a social network analysis of thousands of patients over time requires significant computing resources and is computationally intractable. Because our goal is to identify the coefficient matrix that maximizes the model likelihood, following the assumption for the ERG model that the number of actors and their attributes are fixed, we propose a degenerate statistical model to specify the probability of a set of ties  $Y_t$  at time  $t$  given a set of patients and their individual and pairwise attributes as

$$P(Y_t = y | N \text{ actors}) = \frac{\exp(\sum_k \eta_k Z_{k,t-1}(y))}{1 + \exp(\sum_k \eta_k Z_{k,t-1}(y))}.$$

Specifically, we consider a binary relationship: presence ( $y_{ij} = 1$ ) or absence ( $y_{ij} = 0$ ) where  $y_{ij}$  is defined as the variable for a direct tie established between patient  $i$  (ego) and patient  $j$  (alter). Therefore,  $Y_t$  describes a network with patients as the nodes and ties as their communication (friend) relationships. The model covariates are represented by  $Z_k(y)$ . The vector  $\eta$  contains coefficients that determine the impact of these covariates for a given network and its corresponding features; these are unknown parameters to be estimated.

To be clear, our model assumes that there is no unobserved homophily beyond that on the observable variables. The time-lagged dependent variables (lagged to the prior time period) eliminate serial correlation in the errors when there are more time periods observed (Christakis and Fowler 2013). In particular, we test for significant serial correlation in the error terms using a Lagrange multiplier test. Our results show that the correlation ceases to be significant with a single lagged dependent variable (AR(1) case). We also include lagged social tie  $y_{ij}$  in the model, which helps to control for a patient's endowment or intrinsic predilection to evince a social tie (Christakis and Fowler 2013).

Further, the network among patients is not static. Over time, patients find new members to with whom they can communicate and establish new relationships as well as remove unnecessary ties to reduce the size of the network and the volume of information to be processed. In this way, the configuration of the network changes as patients gradually develop their information management skills. Over a given period of time, as ties are added and removed, the network will contain different configurations. Thus, given an observed configuration of network  $Y_t$  at time  $t$ , the

probability of observing the network distribution for  $T$  periods is given by

$$P(Y_1 Y_2 \cdots Y_T | N \text{ actors}) = \prod_t P(Y_t = y | N \text{ actors}).$$

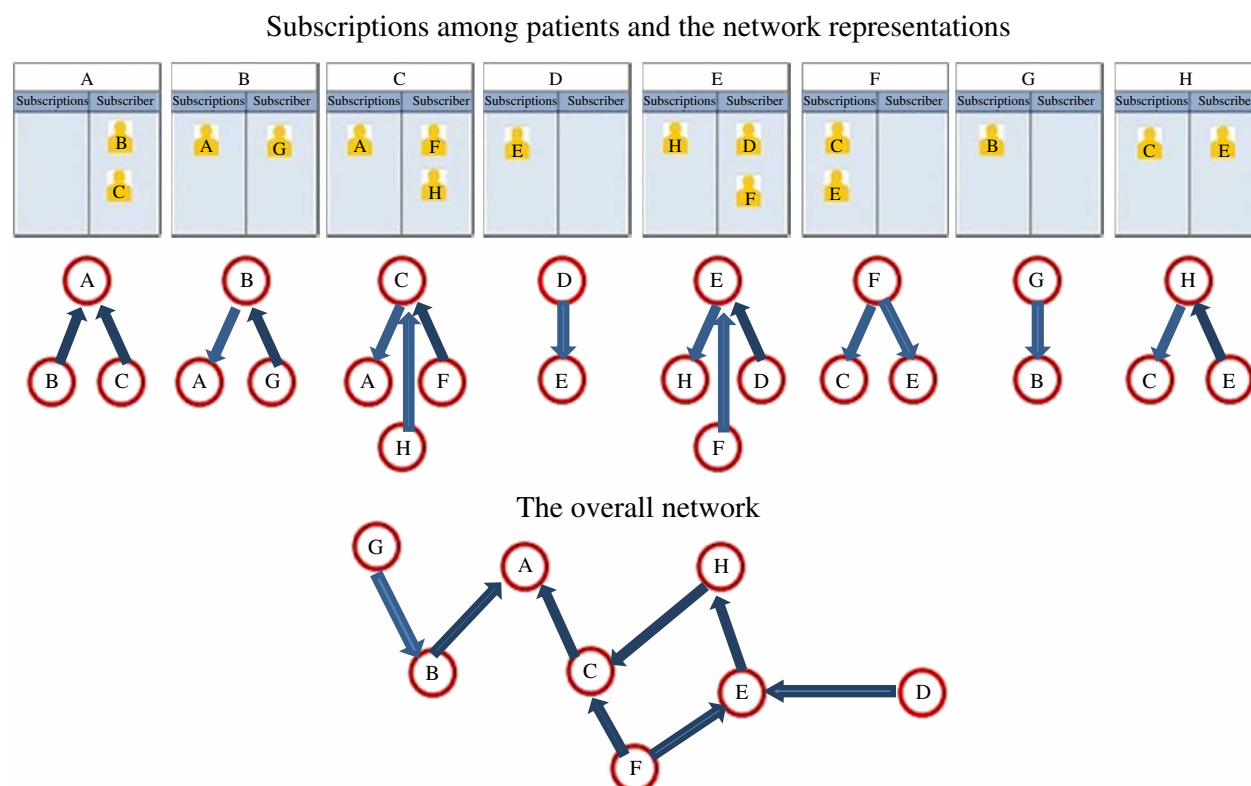
In this actor-driven model, we assume that each patient has control over her outgoing ties  $y_{ij}$  and her characteristics  $Z_k(y)$ .

## 4. Data and Variables

The network considered in this study changed over time as patients progressively expanded their social relations with the exogenous incentives within their cognitive limits. Specifically, the social network from which we collected data is a patients' subscription network. It is easy to create a subscription tie in most online social networking sites. When a patient finds a piece of information that is helpful and wants to keep track of it, there is an icon next to the content, regardless of whether it is a patient's profile or a forum discussion, thereby making it easy for community members to subscribe to it. To remove such a subscription is also straightforward, because there is an "unsubscribe" icon next to the subscribed content.

The subscription feature helps patients to establish a direct social network and exchange information more easily, thereby reducing the needed cognitive effort. On each patient's profile, there are subscription and subscriber tabs that provide patients with an immediate view of the patients to whom she subscribes and those who subscribe to her. On her subscription list, the patient receives notice of updates on patients whom she cares about. The patient also has access to a list of other patients who subscribe to the same patients she does. Figure 2 shows an example of how the social network is constructed. For instance, if patient B subscribes to patient A's profile, on patient B's Web page, there is information about patient A's most recent activities. Others who have connections to patient A would potentially have common interests or concerns similar to those of patient B. Therefore, through the subscriber list of patient A, it is easy for patient B to find similar patients (e.g., patient C) or to look for the information that she needs. At the same time, patient B also can see who has subscribed to her (e.g., patient G) and review each of those patients' information to determine whether it is beneficial to subscribe to the patient's content and, thus, create reciprocal relations.

We collected data for four months, starting in late 2009, from a community focused on mental problems. We chose this type of community because mental disease is a social disease (Coiera 2013), and patients with mental problems need more social support and are more active than are members in other communities. Note that the subscription network is the result of patients' selections of similar patients given

**Figure 2** (Color online) The Construction of a Patient's Social Network

their network structure and cognitive efforts. This endorsement is a one-sided action, and patients can unsubscribe from others' profiles at any time. Under such circumstances, the social network is a directed graph, and the existence or absence of a tie indicates the social relations among patients, whose connection relations change over time. Although the website provides a feature that allows users to remove subscriptions, the implications of deleting can discourage users from doing so (Golbeck 2007).

In our study, we observe that the network changed with a much slower rate of disconnection than relation formation, which is consistent with prior findings. This observation can be further explained with our data that show that most individuals have social relations within their cognitive abilities. At the end of the four-month period, our data set contained the subscription relationship records for 3,708 patients and their relations. Because of computational complexity, we constructed our major networks by focusing on 1,322 patients whose dyadic records were first observed at the beginning of our data collection. Our study covers three time stamps (network layouts). A summary of network statistics is shown in Table 1.

#### 4.1. Patient's Health Profile and Experientially Similar Traits

Typically, a patient's profile contains her basic personal information, such as age, gender, location, city,

**Table 1** Net Change in Dynamics of Social Networks

	Profile subscribers	Network	Edge	Removed	Added	Network density
Network one	1,322	1,322	3,041	—	—	0.0017
Network two	2,517	1,322	8,426	807	6,192	0.0048
Network three	3,708	1,322	19,107	2,082	12,763	0.0109

	Indegree					Outdegree				
	Edge	Mean	Std dev	Min	Max	Mean	Std dev	Min	Max	
Network one	3,041	2.3003	2.1269	0	13	2.3003	2.0750	0	11	
Network two	8,426	6.3722	3.2801	0	18	6.3722	3.2568	0	19	
Network three	19,107	14.4516	4.4833	2	30	14.4516	4.4962	2	29	

and state. Based on her disease category, a patient also can describe her health, illnesses, and other health-related experiences. This voluntarily shared data contains such information as a patient's primary condition, the time of diagnosis, detailed symptoms, and medical and supplementary treatment history. In addition, similar to information contained in other popular online social networking websites, a patient's online activities are also recorded on their profile, including her membership as number of days in the community and the time of her most recent online update. Because we are interested in patients' interactions at the dyadic level, we need to be able to compare characteristics of a pair of patients. Although we believe that experiential proximity is more attractive,

we still track patients' demographic and geographic information. We collect and construct variables based on patients' disease experience-based similarities and the information load for patients to process.

*Demographic and geographic homophily.* We develop the basic homophily measures for demographic and geographic similarity. There are three covariates: *age*, *gender*, and *state*. The first variable is the age (absolute) difference between a patient and her alter, and the latter two are binary variables that indicate whether two patients have the same gender or are from the same geographic location.

*Disease experience-based homophily.* We further construct four disease experience-related variables to assess the experiential similarity and describe the information and knowledge exchange among patients based on their communication ties. Considering a directed communication connection, a pair of patients who established a connection may have been diagnosed with the same disease (*disease* = 1) or have similar firsthand experience with a certain number of symptoms (*symptom*). The covariates *drug experience* and *treatment experience* are count variables that indicate the number of prescription drugs and other treatments that these two patients have taken that are different while they suffer from the same symptoms. Because depression is a disease that has no certain cures and for which no medical treatment has been found to be universally functional among patients, we use the different experiences to represent the best interests of patients, which is to explore more possibilities and possible outcomes.

*Information load.* In the process of comparing and selecting, patients need to process information and experiences related to prescribed drugs and treatments that might be beneficial to them. Given the health condition, the more medical treatments that a patient has, the more information with which she needs to contend. In light of this, we create a covariate *info. load* based on the patient's disease condition. It is constructed by including the number of prescription drugs that the patient has reported using and the number of other methods that she uses in addition to prescription drugs. This variable can be used to measure the amount of health experiential information a patient needs to process as a means to better understand her health condition, improve her disease knowledge, and compare her disease with similar others, which is related to the patient's cognitive effort spent in the expansion of social relations.

#### 4.2. Patients' Network Embeddedness and Cognitive Limits

On each patient's profile, there is a tab associated with her subscription list that helps the patient to construct and manage her friend network. By using this

tool, a patient can subscribe to the patient profiles and forums that interest them. Shared interests can include disease history (which contains useful information for others to learn from), disease treatments, disease progress (which gives firsthand experience of individual outcomes), and discussion topics that deliver meaningful information. The profile owner, the ego, also can see a list of the patients who subscribe to her.

*Network embeddedness.* Network theory assumes that, in information diffusion, the pattern of one's ties is more important than the size of one's network (Burt 2004). A patient who holds a more central position in the network will have greater access to information and resources exchanged from other parts of the network. In particular, researchers have shown that individuals with many structural holes in their ego network obtain a critical position by bridging other sectors of the network and have access to nonredundant information (Burt 2004). At the individual level, the concept of structural holes concerns the lack of connection and social capital that are shown to be accrued to individuals who are connected to others who are not themselves connected (Burt 2004). Following prior research, we use betweenness centrality to measure patients' *network embeddedness* (Sasidharan et al. 2012). A high value on this measure implies a central position in the network and grants an individual the ability to reach various resources.

Other than betweenness centrality, which ensures that patients receive nonredundant information, there are other measures that capture different aspects of network characteristics. For instance, a cohesive or strongly connected subgroup ensures the consistency of information. Intuition suggests that patients in a cluster are connecting more intensely than are others, and such dense relations make the same piece of information transfer from different patients. This ensures the accuracy of the information, although it increases the information redundancy. To capture this aspect, we calculate the clustering coefficient to each patient in the network and create a variable, *dense relation*, to derive the probability that a patient's subscribers also subscribe to each other. Specifically, this measurement describes the relatively strong internal relations shared by a group of people and is calculated by the proportion of connections among the patients within their immediate alters (Wasserman and Faust 1994). Further, the number of two-way connections, *reciprocity*, also suggests redundancy in the information flow. The measure *closeness* centrality focuses on how close an individual is to other patients in a network. It is typically defined as a function of geodesic distances, including direct and indirect connections. A high value of this measure implies a high engagement of a patient in the network and that she



can be involved quickly in communication with all other members of the community.

**Cognitive limits.** The factors associated with cognitive capabilities considered in this study include the network size and cognitive effort for network growth. Because a subscription is a directional connection, we calculate a patient's outdegree in the network, indicating the outgoing connections and social relations started by her, as the measure for a patient's *network size*. Typically, degree centrality is an assessment of the number of immediate connections possessed by an individual and captures the level of social capital immediately available to the ego (Wasserman and Faust 1994). In our context, a patient's outgoing ties suggest her active actions for social support, and this self-motivated activity is restrained by her cognitive capability. That is, the *network size* is constructed by counting the number of outgoing connections of the focal patient.

We calculate the shortest path between a pair of patients as the *network cost* to proxy the cost for them to participate in collaborative health education. Patients need to invest time and effort to reach out for information or knowledge that is not available from their direct connections, and this cost increases with the social distance that they traverse to locate a similar other. When a pair of patients is further apart in the network, the quality and relevance of information and the sense of comfort and trust are likely to deteriorate, hence adding to the cost. In addition, the information load to be processed increases along with the connection path such that patients encounter the problem of information overload with the distance of connections, which can be considered another negative impact of network connections. In our study, the shortest path between patients is constructed as the number of connections for a pair of patients at the minimum distance. If a pair of patients is not connected, the shortest path between those two nodes was set at infinity. To account for this situation, we invert *network cost* and adopt  $\text{network cost}^{-1}$ , or inverse shortest path, in our estimation model.

Because search tools, such as collaborative filtering, can help patients to reduce the time or effort needed to navigate desired information, we construct *search cost* by measuring a patient's posting activities to capture the benefit brought by search tools. The more posts that a patient contributes to the online health-care community, the easier that she can be retrieved through the search feature.

#### 4.3. Dependent Variable and Other Controls

As discussed earlier, individuals in the online health-care community are searching for similar patients and shared information and experiences. They will construct a subscription tie to others only if those health

profiles are informative and/or contain some similarities to their own situation. Under such circumstances, the initiation of a subscription tie from patient  $i$  to patient  $j$  indicates patient  $i$ 's belief in some similarities. In other words, the subscription network is an indication of perceived similarity and communication usefulness. This endorsement is a one-sided action, and patients can unsubscribe from others' profiles at any time. Hence, the social network we consider is a directed graph, and the probability of a relational tie may increase as the characteristics of two individuals become more similar. In our model, the absence or presence of a tie was our dependent variable.

To ensure causal relations in our model, we construct three variables to capture patients' characteristics at the individual level. First, the quality of the personal disease information, *profile quality*, is constructed as a binary variable. That is, it is set to 1 if a patient provides sufficient health information, for example, updating her disease progress continuously to give herself and other patients a useful overview of her health condition, or 0 otherwise. The variable *membership* is also constructed to control for a patient's online experience. It is the log-transformed number of days that the patient has participated in the community. Another variable, *post*, is the patient's online activity control. We calculate the number of posts that a patient contributes to the community, which indicates her attitude toward the online health-care community. She could be an active participant if she is self-motivated and engages in disease management. If the patient is merely lurking within the community, her online activities also will affect her social network construction. Therefore, we specifically control the patient's attitude-specific characteristics toward the consumption of health information in our analysis. Table 2 provides a summary of the variables, their descriptions, and their statistics.

## 5. Results and Findings

There are 5,239,086 observations of subscription ties in our study of patients' network dynamics in the three observation periods. We used a maximum likelihood procedure to estimate the coefficients to best fit our model in describing the network and its changes. Following related literature, we adopted the assumption of homogeneous classes of dyads and took a parsimonious approach. We assumed that each instance was equiprobable (Goodreau et al. 2009); that is, a covariate's effect was weighted the same for all observations. To account for the potential heterogeneities associated with patient attributes, we controlled for the unobserved individual attribute-specific configurations in our estimation procedure. In other words, we further developed an instrumental variable using



**Table 2** Data Description and Statistics

Variable	Description	Mean	Std dev	Min	Max
<i>age</i>	The absolute value of age difference between a pair of patients	14.3586	10.1597	0	42
<i>gender</i>	1 if a pair of patients have the same gender; 0 otherwise	0.4991	0.5000	0	1
<i>state</i>	1 if a pair of patients are from the same state; 0 otherwise	0.5005	0.5000	0	1
<i>disease</i>	1 if a pair of patients are diagnosed in the same category of disease; 0 otherwise	0.5003	0.5000	0	1
<i>symptom</i>	The number of same symptoms a pair of patients suffer	2.4984	1.7071	0	5
<i>drug experience</i>	The number of different prescription drugs a pair of patients take if they suffer one or more of the same symptoms	3.0733	3.7578	0	11
<i>treatment experience</i>	The number of different treatments a pair of patients undergo if they suffer one or more of the same symptoms	2.9565	3.7323	0	11
<i>info. load</i>	The total number of prescription drugs and treatments a patient takes	13.4622	6.1447	2	32
<i>network size</i>	A patient's normalized outgoing connections	0.0058	0.0045	0	0.0189
<i>network embeddedness</i>	A patient's normalized betweenness centrality	0.0020	0.0022	0	0.0208
<i>network cost</i> <sup>-1</sup>	The inverse shortest path between a pair of patients	0.2357	0.1439	0	1
<i>search cost</i> <sup>-1</sup>	The number of posts the alter of a patient (ego) has made <sup>a</sup>	14.6601	8.0937	0	40
<i>profile quality</i>	1 if a patient's profile is awarded with good quality stamp; 0 otherwise	0.1581	0.3648	0	1
<i>membership</i>	The log-transformed membership days	4.9859	0.4915	3.3322	5.7104
<i>posts</i>	The number of posts a patient (ego) has made	14.6601	8.0937	0	40
<i>dense relation</i>	The clustering coefficient considering only 1 neighborhood	0.0126	0.0185	0	0.3333
<i>reciprocity</i>	The number of mutual subscriptions	0.5169	0.7088	0	2
<i>closeness</i>	A patient's normalized closeness centrality	0.3044	0.0932	0	0.4424

<sup>a</sup>The more posts that a patient has contributed to the community, the lower is the *search cost* incurred by other patients to find her. Therefore, *search cost* is measured by the inverse of the number of posts. Similarly, to account for the situation of a zero post, we invert *search cost* and adopt *search cost*<sup>-1</sup> in our estimation model.

a patient's averaged out-degree centrality as a way to capture patients' different online behavior patterns. We explicitly involved the lags of covariates in our model to ensure the causal relationship. We applied the random effect (RE) model because our short panel data set had an extremely large number of entities (Torres-Reyna 2007). The lack of the null case of similarity in the fixed effect (FE) model suggests that RE should be used and that the RE model returns a higher likelihood. Serial correlation and heteroskedasticity were explicitly controlled in the estimation. Table 3 presents a summary of our estimated results.

### 5.1. Personal, Demographic, and Health Problem Similarities

Unlike the results from the study of MySpace (Thelwall 2009), our results show significant evidence of age homophily. Because the variable indicates the age difference between a pair of patients, the negative and significant estimate implies that the larger the age difference, the less chance that two patients will communicate. This result is consistent with the findings of Sutor et al. (1995), who suggested that patients find it easier to communicate with others of similar age and that the information exchanges are useful. By contrast, we find no evidence of state and gender homophily, confirming the diminishing explanatory power of demographic and geographic homophily in the online healthcare networking environment.

As expected, we find that experiential-based similarity (i.e., similarities in disease or symptoms) is more important than value similarity (i.e., similarity in age, gender, or state). Because we restricted our attention to an online community associated with mental problems, members were, broadly speaking, all similar. Further classification or categorization does not provide any additional useful information to patients in the community. Therefore, these similarities, measured by *disease* and *symptom*, *disease* × *symptom*, are found to have no impact on the probability that patients would form connections. However, the shared “true” experience of treating the symptoms with prescription drugs becomes attractive and plays a significant role in patients' communication network.

The positive and significant estimate for *drug experience* suggests a higher probability that patients will communicate if they share the same symptoms while taking different prescription drugs. Because patients in this situation can compare their disease progress and gauge the effectiveness of their current medical plan, this learning from others' experiences can potentially contribute to patients' medical decision making. The insignificant result for *treatment experience* indicates that patients rely primarily on prescription drugs to treat their symptoms, whereas other forms of treatment are indirect and supplementary. Overall, using experienced-based similarity in a patient's selection process appears to be the dominant strategy. Thus, Hypothesis 1 is supported.

**Table 3** Estimation Results<sup>a, b</sup>

Variable	Hypotheses	Coefficient
Disease experience-based similarity		
age		−0.0003*** (0.0001)
gender		−0.0009 (0.0024)
state		0.0027 (0.0024)
disease		0.0027 (0.0043)
symptom		0.0002 (0.0010)
disease × symptom		0.0001 (0.0014)
drug experience	H1(+)	0.0011* (0.0006)
treatment experience	H1(+)	0.0009 (0.0006)
Cognitive capabilities		
network size	H2A	0.3071*** (0.0211)
network size <sup>2</sup>	H2A	−0.5064*** (0.0405)
info. load	H2B(−)	−0.0042*** (0.0002)
drug experience × info. load	H2B(−)	−0.0006* (0.0003)
treatment experience × info. load	H2B(−)	−0.0001 (0.0003)
network embeddedness	H3(−)	−0.9799*** (0.0093)
network cost <sup>−1</sup>	H4(−)	0.9648*** (0.0211)
search cost <sup>−1</sup>		0.0066*** (0.0003)
Individual control		
profile quality		−0.0538*** (0.0033)
posts		0.0030*** (0.0002)
membership		0.0470*** (0.0031)
dense relation		−0.3116*** (0.0557)
reciprocity		−0.0069*** (0.0018)
closeness		8.2061*** (0.0417)
individual heterogeneity control		Yes
lagged social ties		Yes
Constant		0.7756*** (0.0185)

<sup>a</sup> network size, network cost<sup>−1</sup>, and closeness are mean centered. The average VIF is 2.93, and individual VIF is less than 8.79.

<sup>b</sup> network size and network embeddedness are scaled up by 100, (drug experience × info. load) and (treatment experience × info. load) are scaled down by 10.

\* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$ .

## 5.2. Cognitive Capabilities and Impact on Network Structure

In our study, cognitive capabilities are measured at both the individual egocentric level and supradynamic effects beyond the immediate social relations. As suggested by network theory, individuals' network behavior, in terms of making connections with other experientially similar patients, is bounded by cognitive abilities. In support of Hypothesis 2A, we find a curvilinear relationship between establishing new social ties and network size. The linear term is positively significant, whereas the second-order term is negative. Although social support flows through adjacent network connections, and patients with more ties have greater access to such social resources (Brass 1984, Sasidharan et al. 2012), their network growth increases until their cognitive limit is reached. In support of Hypothesis 2B, the negative and significant coefficients for drug experience × info. load provides evidence for the negative impact brought on by information overload. Despite the knowledge that patients can gain from experientially similar others, this learning effect is moderated if there is too much

information to process. In other words, the amount of information that needs to be processed increases with the number of prescription drugs and treatments that a patient is taking. The more drugs and treatments a patient takes, the bigger the load she has to process, resulting in a greater chance of information overload. This, in turn, reduces a patient's gain from learning from others.

Our result shows strong support for Hypothesis 3. A patient with high betweenness centrality, who is thus more embedded in the network, is less likely to make new connections (a negative significant coefficient for network embeddedness) because her position in the network already allows the patient to acquire information from various resources. Adding to the effectiveness of getting diversified rather than redundant information, the negative and significant results for dense relation and reciprocity likewise confirm that patients are more likely to "reach out" for divergent social resources that are not available in their cohesive subgroup. Hypothesis 4 is also supported. As we expected, the cost to establish and maintain connections limits patients' network expansion. Specifically, the positive and significant coefficient network cost<sup>−1</sup> confirms the negative relationship between the probability of establishing new social ties and the inversed network cost. When network cost<sup>−1</sup> is lower, i.e., network cost becomes larger, there is a lower probability of establishing a new social tie. It is also observed in our results that, as expected, a higher search cost makes a patient more difficult to find and, hence, decreases the probability of establishing social contact. The key findings are summarized in Table 4.

## 5.3. Social Climate: Comfort vs. Effectiveness

Our results also shed light on the current online healthcare environment, specifically on its social climate. Social climate is defined as personality of the environment or social network (Moos and Lemke 1992). Qualities of social climate foster social comparison, competence, and exchange of social support. It is presumed that only in a helpful and protective atmosphere will social supportive behavior occur (Hinson Langford et al. 1997).

Research has found that patients' perception of health digitalization is closely tied to concerns about privacy, security, and perceived usefulness (Tang et al. 2006). Individuals' attitudes toward the online social network environment, in addition to individual characteristics and network endogenous features that affect network formation and evolution, also affect the configurations of their close-knit communication constructions. As they do in the physical world, patients must face a dilemma as they decide whether to share their health stories with strangers or engage in social relations with someone whom they believe trustworthy, based on their perceptions. This is a potentially

**Table 4** Summary of Results

Dimension		Hypothesis	Result
Perception of similarity	H1	<i>The more experientially similar a patient (alter) is compared to the focal patient (ego), the more likely that the ego will form a social connection.</i>	Supported
Cognitive limits and information processing	H2A	<i>A patient's network expansion has a curvilinear (inverted-U) relationship with network size.</i>	Supported
	H2B	<i>The more experiential information that a patient needs to process, the less likely that she will establish new social connections.</i>	Supported
Network embeddedness and the connection costs	H3	<i>The more embedded a patient is in the network, the less likely that she will establish new social relations.</i>	Supported
	H4	<i>The probability of establishing new social relations decreases with a patient's cost to explore the network.</i>	Supported

serious problem for the online patient-centric space, as patients decide how much information to share, which is important to the success of online interactions (Meyerson et al. 1996, Jarvenpaa and Leidner 1998, Piccoli and Ives 2003, Coppola et al. 2004). These concerns relate to previously discussed motivations in the social network. On one hand, patients have the incentive to stay within their connections and to feel safe, comforted, and supported. On the other hand, patients also have the incentive to reach out and interact with the environment. Studies have shown that the Internet is used more as an information source than as a means to trust other individuals (Hesse et al. 2005, Powell and Clarke 2006).

Our results provide possible explanations for the social climate and patients' online behavior. Confronted with the threat of being identified offline and encountering various subsequent difficulties, patients need to make choices on how closely they want to embed themselves in the virtual community. The estimated results for *dense relation* and *reciprocity* are found to be negative and significant, which suggests that patients do not tend to "stay" within their connections to form close-knit communication constructions and to feel comfort. Instead, they prefer to reach out and become more effective. These behaviors may be explained by a lack of trust of patients who are concerned about protecting themselves.

These implications are also supported by the control variables, which focus on the patients' purpose in joining the online healthcare community. Patients can use the social media platform to record their disease progress, that is, provide documentation, or they

can choose to contact patients and exchange knowledge. The more detailed the health information that individuals upload, the better the information quality will be. This process not only helps patients to keep track of their disease progress, but it also provides an integrated picture of a disease history from which other patients can learn and show their physicians to better understand their experiences. Although the benefits are easy to see, there are also drawbacks. The significant but negative effect for *profile quality* indicates that patients who have good health profiles are less likely to explore the website and contact other patients. It is likely that these patients are beyond the information-seeking stage and have transitioned to an information-sharing stage. This is a potential problem; that is, without efficient mechanisms, the online healthcare community may deteriorate into a form of a digital health record repository and lose its ability to generate social capital.

#### 5.4. Robustness Checks

To ensure the robustness of our results, we conducted further analyses with several other network controls. Specifically, we deployed *transitivity* and *structural equivalence* as alternative network measures. Typically, *transitivity* is constructed as the ratio of the number of transitive triads in the existing network over the number of possible transitive triads. It explains the degree of closeness in the network that a patient forms. The other network measure, *structural equivalence*, is calculated by considering both people and connections. Two actors are structurally equivalent if they have the same types of ties to the same (types of) people. Patients who subscribe to similar patients' profiles will have a higher probability of identifying more similar patients from this transmittable subscription relationship. Therefore, it can be used as an alternative variable for measuring cohesiveness. Noteworthy in our study is that we used the *dissimilarity* index to measure structural equivalence. We counted and normalized the total number of uncommon neighbors for a pair of patients. Therefore, in our model, the structural similarity level decreases with this index. Further, we used this measure to control for the influence of nonsimilar patients. Concerns related to patients' positions and their communications also might arise, because people outside the target group of members could have an impact on patients' connecting probability.

In our study, subscription initiation was interpreted as a connection, and neighbors were defined as such. It is possible that patients shared common friends outside the subscription network to which we were restricted and that these relationships might have increased the probability of creating a tie in the "insider network." For example, suppose that two



patients have the same number of friends, but one has more friends outside the study group. If these outsiders have members from the group, the patient with more outside connections occupies a more important position in the network. To account for this possibility, *dissimilarity* was calculated based on the connections from the full subscription network and, thus, captured this indirect impact. Hence, this index was constructed based on a patient's acquaintances both inside and outside the subscription network. Controlling for these effects, however, did not lead to statistically significantly different results.

## 6. Discussion and Conclusions

Given the realization that even weak social connections have influential power and that social media create loosely aggregated coalitions of individuals who share a common interest, a new socioculture that emphasizes patients as consumers emerges (Coiera 2013). Further, social media in healthcare is taking an increasingly important role in patient-centric healthcare systems (Kane et al. 2009, Agarwal et al. 2010, Fichman et al. 2011). The emergence of collaborative web-based patient education has changed patients' role in healthcare, allowing individuals to move from passive recipients to active contributors in their health education and decision making.

Along with the easy access and open space of the virtual world, however, patients face the challenge of identifying the right individuals with whom to communicate, and the online healthcare community faces the challenge of depending on its members' collaborative activities and the effectiveness of their online participation. Unlike previous studies, we examine the characteristics of content creators (not the content itself) and their social connections. That is, our work focuses on the communication network structure and the driving factors behind its configuration dynamics. We propose an econometric framework at the dyadic level to analyze longitudinal data. This approach incorporates relevant network-level measures to investigate individual preferences for social contacts in an effort to empirically explain online social network evolution.

Our results exhibit evidence of what drives the network dynamics from two perspectives: patients' (disease) experience-based similarity as well as cognitive capabilities and network embeddedness. Patients are more likely to contact other patients who have the same health problem but are undergoing different medical treatments. The resulting shared experience extends patients' knowledge and helps them to compare their health condition with that of others. In addition, as a result of the unknown quality and credibility of the information in online healthcare

communities, patients are more likely to depend on existing connections to expand their social networks. Yet, a patient's social relations are still constrained by her cognitive ability as well as her network embeddedness. For instance, the problem of information overload limits the expansion of patients' communication. The cost to establish new social relations and trade-offs between the cost and gain also limit patients' network growth. Concerns about sharing health information also bring a negative aspect that can shape patients' online activities. These findings are of special interest to the healthcare industry, which needs to better engage with online social networking websites and improve their services.

The reason that online healthcare communities attract so much attention is that they create the opportunity for patients to work collaboratively, which affects medical disciplines such as evidence-based and personalized medicine. Patients are more likely to share their personal health information when they believe that it can help others (Anderson and Agarwal 2011), and, thus, they voluntarily contribute their time, effort, and knowledge to help achieve collaborative benefits (Wasko and Faraj 2005). These potential contributors are diverse in knowledge and motivation, and determining how to increase the value of knowledge sharing and creation is the challenge that faces these new social media platforms, which is largely reflected in the network constructed in these communities.

Our results indicate that patients are unlikely to form close connections (or cohesive groups) in online healthcare communities, as shown by the significant but negative coefficient for *dense relation*. Because a cohesive subgroup indicates the existence of redundant links, patients are more attracted to the different experiences and knowledge that are typically brought by outliers. This is similar to the difference between strong and weak ties, whereas a strong tie ensures the speed and effectiveness of communication, a weak tie always brings in new characteristics and ideas. This observation supports the conjecture that the purpose of using subscription tools is to reduce the problem of information overload at the cost of the available resource. All of this suggests that patients focus more on the diversity of information over concern for reliability, suggesting the need for the website to enhance the service and improve information quality. Various mechanisms should be deployed to enhance patients' desire to give and receive information. Although the online media makes it difficult to control for information quality and knowledge exchange, a mechanism for fostering trust would be a good initial step because trust is strongly related to information disclosure (Ridings et al. 2002, Metzger 2004).



Participation in an online healthcare community is a novel way to make personal health records accessible for community members. This community technology is believed to contribute to the transformation of healthcare delivery systems. Future research should further investigate customized search engines to deliver the most relevant and timely information. New recommendation mechanisms should be relationship based and designed by taking network characteristics and other context features (e.g., privacy concerns about health data) into consideration. Another promising extension could be a comparison of the network dynamics between the healthcare environment and the nonhealthcare setting. The driving forces that we identified in this study may or may not apply to other areas. The third possible extension is related to the objectives of patients' online activities. For example, patients may join online healthcare communities to socialize with other patients, or patients may look for relevant information and other patients' experiences. These different objectives should lead to distinct behavior patterns. Understanding what patients need will help practitioners to provide better service and to meet patients' requirements. Although these important questions are difficult to address in our study setting, we believe they could be studied experimentally.

## Acknowledgments

The authors thank the senior editor, associate editor, and the three anonymous reviewers for their constructive suggestions throughout the review process. The third author was supported in part by the National Natural Science Foundation of China [Grant 71328103].

## References

- Agarwal R, Gao G, DesRoches C, Jha AK (2010) The digital transformation of healthcare: Current status and the road ahead. *Inform. Systems Res.* 21(4):796–809.
- Anderson CL, Agarwal R (2011) The digitization of healthcare: Boundary risks, emotion, and consumer willingness to disclose personal health information. *Inform. Systems Res.* 21(4):796–809.
- Barrera M (1986) Distinctions between social support concepts, measures, and models. *Amer. J. Community Psych.* 14(4):413–422.
- Bawden D, Robinson L (2009) The dark side of information: Overload, anxiety and other paradoxes and pathologies. *J. Inform. Sci.* 35(2):180–191.
- Berkman LF, Glass T (2000) Social integration, social networks, social support, and health. Berkman LF, Kawachi I, eds. *Social Epidemiology* (Oxford University Press, New York), 137–173.
- Berkman LF, Syme SL (1979) Social networks, host-resistance, and mortality—9-year follow-up study of Alameda County residents. *Amer. J. Epidemiology* 109(2):186–204.
- Borgatti SP, Mehra A, Brass D, Labianca G (2009) Network analysis in the social sciences. *Science* 323(5916):892–895.
- Brass D (1984) Being in the right place: A structural analysis of individual influence in an organization. *Admin. Sci. Quart.* 29(4):518–539.
- Bundorf MK, Wagner TH, Singer SJ, Baker LC (2006) Who searches the Internet for health information? *Health Services Res.* 41(3 Pt 1):819–836.
- Burt RS (2004) Structural holes and good ideas. *Amer. J. Sociol.* 110(2):349–399.
- Byrne D (1961) Interpersonal attraction and attitude similarity. *J. Abnormal Soc. Psych.* 62(3):713–715.
- Byrne D (1971) *The Attraction Paradigm* (Academic Press, Orlando, FL).
- Cassel J (1976) Contribution of social-environment to host-resistance—4th Wade Hampton Frost Lecture. *Amer. J. Epidemiology* 140(2):107–123.
- Christakis NA, Fowler JH (2013) Social contagion theory: Examining dynamic social networks and human behavior. *Statist. Med.* 32(4):556–577.
- Christensen MH, Boisse N, Sanchez W, Friedmann PD (2004) Enhancing vocational rehabilitation counselors' substance abuse screening and brief intervention practices. *J. Vocational Rehabilitation* 21(3):157–163.
- Cline RJW, Haynes KM (2001) Consumer health information seeking on the Internet: The state of the art. *Health Educ. Res.* 16(6):671–692.
- Cobb S (1976) Social support as a moderator of life stress. *Psychosom. Med.* 38(5):300–314.
- Coiera E (2013) Social networks, social media, and social diseases. *BMJ* 346:f3007.
- Coppola N, Hiltz SR, Rotter N (2004) Building trust in virtual teams. *IEEE Trans. Professional Comm.* 47(2):95–104.
- Denning P (2002) Internet time out. *Comm. ACM* 45(33):15–18.
- Edmunds A, Morris A (2000) The problem of information overload in business organizations: A review of the literature. *Internat. J. Inform. Management* 20(1):17–28.
- Eysenbach G (2008) Medicine 2.0: Social networking, collaboration, participation, apomediation, and openness. *J. Med. Internet Res.* 10(3):e22.
- Fichman RG, Kohli R, Krishnan R, Kane GC (2011) The role of information systems in healthcare: Current research and future trends. *Inform. Systems Res.* 22(3):419–428.
- Frost JH, Massagli MP (2008) Social uses of personal health information within PatientsLikeMe, an online patient community: What can happen when patients have access to one another's data. *J. Med. Internet Res.* 10(3):e15.
- Gage EA (2013) Social networks of experientially similar others: Formation, activation, and consequences of network ties on the health care experience. *Soc. Sci. Medicine* 95:43–51.
- Golbeck J (2007) The dynamics of web-based social networks: Membership, relationships, and change. *First Monday* 12(11).
- Goodreau SM, Kitts JA, Morris M (2009) Birds of a feature, or friend of a friend? Using exponential random graph models to investigate adolescent social networks. *Demography* 46(1):103–125.
- Greenberg J (1991) *Oedipus and Beyond: A Clinical Theory* (Harvard University Press, Cambridge, MA).
- Grewal R, Lilien GL, Mallapragada G (2006) Location, location, location: How network embeddedness affects project success in open source systems. *Management Sci.* 52(7):1043–1056.
- Gross BM (1964) *The Managing of Organizations: The Administrative Struggle* (Free Press, New York).
- Gruhl D, Meredith DN, Pieper JH, Cozzi A, Dill S (2006) The Web beyond popularity. *Proc. 15th Internat. World Wide Web Conf., WWW '06* (ACM Press, New York), 183–192.
- Hesse BW, Nelson DE, Kreps GL, Croyle RT, Arora NK, Rimer BK, Viswanath R (2005) Trust and sources of health information: the impact of the Internet and its implications for health care providers: Findings from the first health information national trends survey. *Arch Intern. Med.* 165(22):2618–2624.
- Hinson Langford CP, Bowsher J, Moloney JP, Lillis PP (1997) Social support: A conceptual analysis. *J. Advanced Nursing* 25(1):95–100.
- Jarvenpaa S, Leidner D (1998) Communication and trust in global virtual teams. *Organ. Sci.* 10(6):791–815.

- Jones Q, Ravid G, Rafaeli S (2004) Information overload and the message dynamics of online interaction spaces: A theoretical model and empirical exploration. *Inform. Systems Res.* 15(2):194–210.
- Kadushin C (2012) *Understanding Social Networks: Theories, Concepts, and Findings* (Oxford University Press, Oxford, UK).
- Kane GC (2011) A multi-method study of information quality in wiki collaboration. *ACM Trans. Management Inform. Systems* 2(1):1–16.
- Kane GC, Fichman RG, Gallaughier J, Glaser J (2009) Community relations 2.0. *Harvard Bus. Rev.* 87(11):45–50.
- Lin N (2001) *Social Capital: A Theory of Structure and Action* (Cambridge University Press, London).
- McPherson M, Smith-Lovin L, Brashears ME (2006) Social isolation in America: Changes in core discussion networks over two decades. *Amer. Sociol. Rev.* 71(3):353–375.
- McPherson M, Smith-Lovin L, Cook JM (2001) Birds of a feather: Homophily in social networks. *Annual Rev. Sociol.* 27:415–444.
- Metzger MJ (2004) Privacy, trust, and disclosure: Exploring barriers to electronic commerce. *J. Comput.-Mediated Comm.* 9(4).
- Meyerson D, Weick K, Kramer RM (1996) Swift trust and temporary groups. *Trust in Organizations: Frontiers of Theory and Research* (Sage Publications, Thousand Oaks, CA).
- Monge PR, Contractor N (2003) *Theories of Communication Networks* (Oxford University Press, Oxford, UK).
- Moos R, Lemke S (1992) Sheltered care environment scale manual. Department of Veterans Affairs and Stanford University Medical Centers, Palo Alto, CA.
- Nahapiet J, Ghoshal S (1998) Social capital, intellectual capital and the organizational advantage. *Acad. Management Rev.* 23(2):242–266.
- O'Grady LA, Witteman H, Wathen CN (2008) The experiential health information processing model: Supporting collaborative web-based patient education. *BMC Medical Informatics Decision Making* 8(58).
- Parks MR (2007) *Personal Networks and Personal Relationships* (Lawrence Erlbaum, Mahwah, NJ).
- Piccoli G, Ives B (2003) Trust and the unintended effects of behavior control in virtual teams. *MIS Quart.* 27(3):365–395.
- Powell J, Clarke A (2006) Internet information-seeking in mental health: Population survey. *British J. Psychiatry* 189:273–277.
- Preece J, Ghazati K (1998) In search of empathy online: A review of 100 online communities. *Proc. Assoc. Inform. Systems Americas Conf.*, 92–94.
- Preece J, Ghazati K (2001) Observations and explorations of empathy online. Rice RR, Katz JE, eds. *The Internet and Health Communication: Experience and Expectations* (Sage Publications, Thousand Oaks, CA), 237–260.
- Putnam RD (1995) The prosperous community: Social capital and public life. *Amer. Prospect.* 13:35–42.
- PwC HRI Social Media Consumer Survey (2012) Social media “likes” healthcare: From marketing to social business. <http://www.pwc.com/us/en/health-industries/publications/health-care-social-media.jhtml>.
- Ransbotham S, Kane G, Lurie N (2012) Network characteristics and the value of collaborative user-generated content. *Marketing Sci.* 31(3):387–405.
- Reagans R (2005) Preferences, identity, and competition: Predicting tie strength from demographic data. *Management Sci.* 51(9):1374–1383.
- Ridings CM, Gefen D, Arinze B (2002) Some antecedents and effects of trust in virtual communities. *J. Strategic Inform. Systems* 11(3–4):271–295.
- Sasidharan S, Santhanam R, Brass DJ, Sambamurthy V (2012) The effects of social network structure on enterprise systems success: A longitudinal multilevel analysis. *Inform. Systems Res.* 23(3–1):658–678.
- Scheffler RM, Brown TT, Syme L, Kawachi I, Tolstyk I, Iribarren C (2008) Community-level social capital and recurrence of acute coronary syndrome. *Soc. Sci. Medicine* 66(7):1603–1613.
- Shumaker SA, Brownell A (1984) Toward a theory of social support: Closing conceptual gaps. *J. Soc. Issues* 40(4):11–33.
- Smith KP, Christakis NA (2008) Social networks and health. *Annu. Rev. Sociol.* 34:405–429.
- Snijders TAB (2001) The statistical evaluation of social network dynamics. Sobel ME, Becker MP, eds. *Sociological Methodology* (Basil Blackwell, Boston), 361–395.
- Snijders TAB (2005) Models for longitudinal network data. Carrington P, Scott J, Wasserman S, eds. *Models and Methods in Social Network Analysis* (Cambridge University Press, New York), 215–247.
- Stewart MJ (1993) *Integrating Social Support in Nursing* (Sage Publications, Newbury Park, CA).
- Suitor JJ, Pillemer K, Keeton S (1995) When experience counts—The effects of experiential and structural similarity of patterns of support and interpersonal stress. *Soc. Forces* 73(4):1573–1588.
- Swan M (2009) Emerging patient-driven health care models: An examination of health social networks, consumer personalized medicine and quantified self-tracking. *Internat. J. Environ. Res. Public Health* 6(2):492–525.
- Swann WB, Brown JD (1990) From self to health: Self-verification and identity disruption. Sarason BR, Sarason IG, Pierce GE, eds. *Social Support: An Interactional View* (John Wiley and Sons, New York), 150–172.
- Tang PC, Ash JS, Bates DW, Overhage JM, Sands DZ (2006) Personal health record: Definitions, benefits, and strategies for overcoming barriers to adoption. *J. Amer. Medical Informatics Assoc.* 13(2):121–126.
- Thelwall M (2009) Homophily in mySpace. *J. Amer. Soc. Inform. Sci. Tech.* 60(2):219–231.
- Thoits PA (2010) Stress and health: Major findings and policy implications. *J. Health Soc. Behav.* 51:S41–S53.
- Thoits PA (2011) Mechanisms linking social ties and support to physical and mental health. *J. Health Soc. Behav.* 35:53–79.
- Torres-Reyna O (2007) Panel data analysis fixed and random effects using Stata (v.4.2). <http://www.princeton.edu/~otorres/Panel101.pdf>.
- Uphoff H (2000) Understanding social capital: Learning from the analysis and experience of participation. Dasgupta P, Serageldin I, eds. *Social Capital: A Multifaceted Perspective* (The World Bank, Washington, DC), 215–249.
- Wang ZM, Walther JB, Pingree S, Hawkins RP (2008) Health information, credibility, homophily, and influence via the Internet: Web sites versus discussion groups. *Health Comm.* 23(4):358–368.
- Wasko M, Faraj S (2005) Why should I share? Examine social capital and knowledge contribution in electronic networks of practice. *MIS Quart.* 29(1):35–57.
- Wasserman S, Faust K (1994) *Social Network Analysis: Methods and Application* (Cambridge University Press, New York).
- Weiss LA, Shen Y, Korn JM, Arking DE, Miller DT, Fossdal R, Saemundsen E, Stefansson H, Ferreira MA, Green T, Platt OS, et al. (2008) Association between microdeletion and microduplication at 16p11.2 and autism. *New England J. Medicine* 358(7):667–675.
- Yan L, Tan Y (2014) Feel blue so go online? An empirical study of online supports among patents. *Inform. Systems Res.* 25(4):690–709.
- Yellowlees PM, Brooks PM (1999) Health online: The future isn't what it used to be. *Medical J. Australia* 191:522–525.
- Zetterberg HL (2011) *The Many-Splendored Society. Book 2. A Language-Based Edifice of Social Structures*, 2nd ed.