Chapter 15
Leveraging Social Media for Health Promotion and Behavior Change: Methods of Analysis and Opportunities for Intervention

Sahiti Myneni, Kayo Fujimoto, and Trevor Cohen

Abstract This chapter describes methodologies used to describe, model, and predict user communication patterns in social media interactions, with the shared goal of facilitating understanding of health-related behavior change. To set the stage, the chapter presents an overview of the documented effects of social relationships on health behavior change. Investigators from a variety of disciplines have attempted to understand and harness these social ties for health promotion. Online communities, which digitize peer-to-peer communication, provide a unique opportunity to researchers to understand the mechanisms underlying human behavior change. Through transdisciplinary methods that draw upon socio-behavioral theories, and information and network sciences, analysis of communication patterns underlying social media user interactions is possible at scale. Such methods can provide insight into development of "healthier life" technologies that harness the power of social connections. Examples of such translational projects and implications for public health practice are discussed to conclude the chapter.

Keywords Behavior change • Social media • Online social networks • Semantics • Network analysis • Health promotion • Smoking cessation • Social influence • Behavior change theories

---

S. Myneni, Ph.D., M.S.E. (✉) • T. Cohen, M.B.Ch.B., Ph.D.
School of Biomedical Informatics, University of Texas Health Science Center at Houston, 7000 Fannin St, Suite 165, Houston, TX 77584, USA
e-mail: sahiti.myneni@uth.tmc.edu

K. Fujimoto, Ph.D.
Division of Health Promotion and Behavioral Sciences, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA

V.L. Patel et al. (eds.), *Cognitive Informatics in Health and Biomedicine*, Health Informatics, DOI 10.1007/978-3-319-51732-2_15

15.1 The Role of Social Media in Health Behavior Change

Social media can be defined as the collective forms of electronic communication (e.g. web sites for social networking and microblogging) through which users create online communities to share information, ideas, personal messages, and other content (e.g. videos) (Kaplan and Haenlein 2010; Kietzmann et al. 2011). Healthcare consumers employ these online venues to interact with peers and care providers, to gain knowledge and reciprocate social support on a variety of health-related issues (Chou et al. 2009; Hawn 2009; Sarasohn-Kahn 2008; Fisher and Clayton 2012). From a research and intervention perspective, the advantages of such platforms over traditional pamphlet-based and web-based static health education material include user-generated online content; wide reach and just-in-time accessibility via web and mobile connectivity; the ability to involve multiple stakeholders (e.g. public health agencies) in the content dissemination process, and research potential to better understand participants' health-related needs (e.g. behavioral, tactical, knowledge-based) (Korda and Itani 2013). On account of these advantages, social media platforms have the potential to facilitate interventions and studies involving large numbers of participants, without undue expense. Interventions disseminated on such platforms have the potential for great impact, on account of the capability to leverage *social influence*.

The influence of social relationships and social support on health behaviors has been well documented (Heaney and Israel 2008; Umberson and Montez 2010). While associations between social relationships and health are complex and not necessarily causal in nature (Lyons 2011), evidence suggests that the positive health enhancing effects of social relationships can be used for the promotion of healthy behaviors such as smoking cessation, physical activity, and nutrition management (McGloin and Eslami 2015; Cavallo et al. 2012, 2014; Cobb et al. 2005). Christakis and Fowler's influential analysis of the Framingham dataset shows an association between the smoking behavior within an individual's social network and the likelihood that this individual will quit smoking (Christakis and Fowler 2008). Similar observations have been made with other psychosocial behaviors (Christakis and Fowler 2007, 2008; Fowler and Christakis 2008; Rosenquist et al. 2010). For example, smoking cessation by a spouse, sibling or friend decreased a subject's chance of smoking by 67%, 25% and 36% respectively, as compared with having no social contacts attempting to quit. In contrast to these positive effects, peer group studies on smoking behavior revealed that smokers are more likely to have friends who are also smokers (Alexander et al. 2001). Several observational studies have shown the effects of social constructs such as social influence, selection, norms, and consequences on an individual's involvement in risky behaviors (Valente 2010). Studies conducted on social relationships have revealed different types of social influence, referred to as peer influence, affiliation influence, and positional influence. Peer influence is a form of direct influence that is explicitly based on direct friendship relations (Alexander et al. 2001). Positional influence defines a form of influence that is exerted by an individual as a result of their occupying a central position in a social network (Freeman 1978). Affiliation-based influence takes into account indirect sources of influence including participation in organized group activities and events (Fujimoto et al. 2012). In summary, social networks can be viewed as complex ecosystems that can have positive and negative influences on health behaviors.

A number of "offline" community-based interventions have been developed and evaluated for *health promotion* and *behavior change*. Such programs generally attempt to segment a target population in order to tailor messages according to gender, age, culture, and so forth. Examples include the Minnesota (Perry et al. 1992), Pawtucket (Elder et al. 1986), Stanford Heart Health (Killen et al. 1989), and COMMIT programs (Royce et al. 1993). With tremendous uptake of social media platforms, users of these online venues turn to their peers to share experiences, ask questions, provide emotional support, and exchange self-help advice with one another. And, since the communication events on these platforms are electronically captured, it may be possible for researchers to (a) understand and predict the sources of social influence, and resulting effects on health behaviors, and (b) model interventions that apply data-driven analytics of social media interactions to inform the design of targeted user-information interactions.

The next sections of this chapter are organized as follows. Firstly, we provide the current landscape of social media uptake in health care research, and summarize the research and intervention opportunities that are made possible in recent times by virtue of health consumer driven social media platforms. Secondly, we present an overview of methodologies that are theory-driven, quantitative, and semi-automated in nature and describe the ways in which these methods can be used to describe, visualize, and model user interactions in social media platforms. We then discuss the ways in which these methods can lead us to social influence patterns underlying peer-to-peer communication events in social media platforms to explain individual and group level *behavior change*. Thirdly, we summarize a series of studies that describe the application of these methods to understand social media interactions in a health-related online community for *smoking cessation*. Finally, the implications for intervention design and public health practice are discussed.

15.2 Overview of Social Platforms in Digital Era

In recent years, the penetration of online social media into everyday lives has been astonishing. More than a billion people (1/7th of the world's population) now use a single social networking service, Facebook (Facebook 2016). Around 72% of American adults have an active account in a social networking service website such as Facebook, Twitter, Linkedin, and Google+ (Duggan et al. 2015). To curb the growing health care costs and improve the efficiency of health and wellness programs, it may be possible to exploit the advantage of these scalable platforms to positively influence health behaviors of individuals, as these networks have the capability to deliver interventions to large populations. Today's social media platforms can be broadly classified as (1) General-purpose social networks, or (2) Activity-specific social network. General-purpose networks such as Twitter and Facebook support social interactions on any topic, while activity-specific networks such as PatientsLikeMe (2016) and QuitNet (2016) provide platforms for participants seeking targeted interactions pertinent to health-related goals. A variety of socio-behavioral interventions have been developed to support healthy lifestyle changes by facilitating attitudinal change, behavioral adherence, and the availability of a support network (Centola 2013; Tang et al. 2015; Prochaska and Prochaska 2011). On account of the availability and accessibility of the World Wide Web via mobile phones, social network interventions can occur in real-time (e.g. community support via mood sensing using a smartphone (Ahmed et al. 2015; Bachmann 2015)), and their capacity for perpetual data collection can provide a rich documentation of habitual behaviors (such as visiting a neighborhood bar) that may influence behavior change (Cohn et al. 2011; Heron and Smyth 2010; Shiffman et al. 2008). Therefore, online social media platforms form the basis for ecological momentary assessments and interventions. These networks form a core component of Health 2.0, which is defined as "user-generated health care promoting patient empowerment and participation" (Van De Belt et al. 2010). As the networks mature with scale, their social value increases, and their data can provide valuable insights into fundamental questions of human behavior. Studies of such network data provide a desirable alternative to traditional retrospective cohort-based investigations, because of their scale, structural control, measurement, replicability, and behavioral fidelity (Centola 2013).

15.3 Opportunities for Research and Implications for Public Health Practice

These virtual platforms open new and important avenues for research, including extending existing socio-behavioral theories to technology-driven interventions; understanding fundamental mechanisms of behavior change; and formulating and evaluating novel interventional approaches. Are theoretical models of socio-behavioral change developed prior to ubiquitous digital communication applicable to both offline and online contexts? (Riley et al. 2011). Social network data contain traces of the cognition and immediate behavior of a person as they attempt to introduce a new change or sustain an existing behavior modification. This provides researchers with an unprecedented opportunity to refine existing theories and models of social networks, social support, and behavior-change that were formulated based on face-to-face communication. Consequently, there is a pressing need for the formulation of new methods and metrics to capture and analyze data patterns derived from *online social networks* to inform our understanding of behavior change at the individual and network level. These new methodological approaches must scale to large online social network datasets. Many network analysis studies do not consider communication content and focus on network structure (Cobb et al. 2005; Centola 2010; Aral et al. 2009; Shalizi and Thomas 2011; Poirier and Cobb 2012).

In contrast, those studies that have considered content have adopted qualitative methods thus limiting their scalability to larger datasets (Myneni et al. 2016a; Hwang et al. 2010; Zhang et al. 2013). So the development of methods to facilitate the inclusion of content into network models is another important research direction. Most current *behavior change theories* suggest strategies that tailor the content that is delivered as part of an intervention. Understanding and leveraging network content may allow us to develop theoretically-grounded approaches to intervention that operate at a network level, thus enhancing their impact and efficacy. Other interesting potential research strategies include: (a) formulating new methods to identify important nodes within online social networks to tailor or deliver a behavior change intervention, (b) disentangling the sociobehavioral factors underlying communication patterns and user engagement in health behaviors, (c) comparative effectiveness research on interventions to see if findings based on retrospective self-reports with sparse observations in the real world are consistent with those based on behavioral data collected online, and (d) interfacing online social networks with other elements of health care such as physicians, insurance providers, and workplace wellness influencers.

15.4 Understanding Communication in Social Media

15.4.1 Theory-Driven Techniques

Prior to the advent of online social networks, several health behavior theories and models have been formulated to attempt to explain behavior change. These theories and models have served as guides for the development and evaluation of both face-to-face and online interventions. The Health Belief Model (Hochbaum et al. 1952), Theory of Planned Behavior (Ajzen 1985), and the Transtheoretical Model (Prochaska and Velicer 1997) belong to the category of intrapersonal models, while Social Cognitive Theory (Bandura 1986) and Social network, support models (Heaney and Israel 2008) are classified as interpersonal models. Intrapersonal models consider individual characteristics that influence behavior, such as knowledge, attitudes, and beliefs. On the other hand, interpersonal models consider group-level dynamics involving family, friends, peers, that provide social identity and support. In the sections that follow, we will describe the main features of those models and theories that have been used most widely as a basis for the design of behavior change interventions.

Health Belief Model (HBM)

This is one of the most widely used conceptual frameworks for explaining and changing individual health behavior. The HBM evolved from a cognitive theory perspective and is a value-expectancy theory, which attempts to explain and predict individual's attitudes toward objects and actions (Hochbaum et al. 1952).

Major components in the HBM include perceived susceptibility; perceived severity; perceived benefits; perceived barriers; cues to action; and self-efficacy. An individual's perceptions of a behavior can be used as predictors of behavior change outcomes under certain conditions that are dependent on demographic (e.g. age, gender) psychosocial (e.g. personality, social class), and structural variables (e.g. prior knowledge, experience). The HBM has been applied to many important healthcare problems focusing on behavioral adherence such as seat belt use (Fernandes et al. 2010).

The Theory of Reasoned Action (TRA)

The TRA suggests behavior is determined by behavioral intention (Fishbein 1979). The intent of a behavior is a function of the person's attitude toward the behavior, their subjective norm associated with the behavior, and their perceived behavioral control. Application areas for this theory include health-related behavior concerning disease prevention and birth control (Albarracin et al. 2001; Fisher et al. 1995).

Social Cognitive Theory (SCT)

The SCT is a theory based on reciprocal determinism between a behavior, the environment, and a person (Bandura 1986). This theory emphasizes self-efficacy, an important concept related to self-confidence. Self-efficacy is defined as "people's judgments of their capabilities to organize and execute courses of action required to attain designated types of performances" (p. 391). Current literature agrees on a common definition that self-efficacy "refers to what a person believes he or she can do on a particular task" (p. 506). Small-scale goal attainment, and confidence building through self-monitoring and continuous feedback is often used to improve a person's self-efficacy. Other important constructs in SCT include behavioral capability; observational learning; reinforcement; outcome expectations and expectancies; emotional coping and self-control. The construct of 'observational learning' has been used by network scientists to provide an explanation for social influence and network clustering of people engaging in the same health behavior (Bandura 2001, 2011). According to SCT, observational learning in behavior change occurs when an individual observes another person engage in a given behavior and receive reinforcements. Another component of SCT called reciprocal determinism takes into account the interactions among individuals, their environments, and behavior goals. The environment in SCT refers to a conglomeration of factors that are external to the individual including his/her social network—family, friends, and peers, and physical objects that might affect behaviors. In case of smoking the physical objects can include availability of patches, access to smoking-designated areas in the work place, and so forth.

The Transtheoretical Model of Change (TTM)

The TTM tries to explain behavior change mechanisms by synthesizing several constructs drawn from other theories (Prochaska and Velicer 1997). *Stages* and *processes* of change are the two main components of TTM. The former component explores the temporality of behavior change, while the later encompasses cognitive and behavioral concepts such as decisional balance (comparative analysis of pros and cons of a proposed behavior change), self-efficacy, and rewards. According to this model, precontemplation, contemplation, preparation, action, maintenance, and termination are the six stages of change, where each stage involves one or more processes of progress. This theory has been successfully applied in several behavior change settings (Prochaska and Prochaska 2011).

Behavior Change Taxonomy

Abraham *et al.* defined a set of "theory-linked" behavior change techniques that can be used to characterize and differentiate between different types of intervention content (Abraham and Michie 2008; Michie et al. 2013). Their taxonomy of 93 theory-linked techniques is the first step towards creating a model of intervention content in the context of theory-driven behavior change constructs. A single behavior change technique can be related to similar behavior change processes from multiple theories. The taxonomy provides a common vocabulary to understand the ways that sociobehavioral and cognitive constructs of the existing behavior change theories have been operationalized in a specific intervention.

15.4.2 Review of Qualitative Studies

Prior qualitative studies on online community interactions have focused on (1) studies of user perceptions of the utility of online communities for a specific health-related illness (e.g., mental health (Donovan 2014), alcoholism (Chuang and Yang 2012), cancer (Klemm et al. 2003), Huntington's disease (Coulson et al. 2007)); (2) characterization of the general conversational interests of specific population (e.g., the elderly (Nimrod 2010)); (3) identification of social support categories (House et al. 1981, 1985) (e.g., informational support, emotional support). Another type of qualitatively-driven social media study involves researchers identifying themselves as such and gathering information in the form of online semi-structured interviews, online focus groups, or internet based surveys to attempt to understand consumers' use of social media. Hwang et al. conducted a network-based survey on the Sparkpeople forum, where members focus on weight loss regimen. The qualitative survey data were analyzed for social support themes using grounded theory techniques. Results indicated that the major social support themes were encouragement and motivation, information and shared experiences (Hwang et al. 2010). In our own work, we conducted qualitative analysis of messages exchanged in QuitNet, which provided insights into the nature of communication events in this community using a combination of the aforementioned theories and behavior change taxonomy. These ranged from discussions on nicotine replacement therapies to stress management strategies (Myneni et al. 2016a). Detailed discussion of these results is provided in the subsequent sections of this chapter.

15.4.3 Automated Methods of Text Analysis

Recent advances in automated text analysis allow for large-scale analysis of the content of communication between members. In this section we review research that has leveraged automated methods of text analysis in an effort to interpret content produced by members of online social networks, with a focus on research in the area of health-related behavior change. However, before we proceed to domain-specific research, we will review general domain research that covers this methodological territory. A recurring theme has to do with the issue of semantic relatedness, on account of the need to identify connections between messages that are similar in meaning, but may not express this meaning using the same words as one another.

Content Analysis in the General Domain

Semantic analysis of social network content using automated methods has been previously applied to the study of research communities in the field of enterprise interoperability. Velardi et al. performed content-based social *network analysis* with the aid of linguistic analysis, text mining, and clustering techniques, in which the semantic relatedness between terms was measured using a taxonomy-based approach (Velardi et al. 2008). Meta-data based approaches have also been used to derive person-word relations (e.g. author-specialization) by extracting social network information using semantic approaches (Matsuo et al. 2007). Classification of conversational and informational questions on social Q&A websites such as Yahoo! Answers has also been attempted using a combination of human coding, statistical analysis, and machine learning (Harper et al. 2009). Another application area of automated natural language processing method is the development of a consumer health vocabulary based on threaded discussions in online social network websites (Doing-Harris and Zeng-Treitler 2011). Doing-Harris et al. have developed a computer-assisted update (CAU) system that consisted of three main parts: a Web crawler and an HTML parser, a candidate term filter that utilizes natural language processing tools including term recognition methods, and a human review interface. The CAU system was applied to the health-related social network website PatientsLikeMe.com to develop and dynamically update the health vocabulary

(Smith and Wicks 2008). Another avenue for automated methods in analysis of online social media content is assessing similarity between two separate texts to derive and understand content structure. Content similarity has been used as a filtering metric along with link analysis to rank influential users in a web forum (Tang and Yang 2010). A precedent for research employing both network models and estimates of semantic relatedness can be found in the psychological literature. For example, Pathfinder networks (PFNETS) employ a scaling technique that builds on relatedness between nodes. If each node represents a concept, the weights of links (or edges) present in the network are defined using human estimates of the relatedness between all pairs of concepts (Schvaneveldt 1990). Consequently the structure of a PFNET is determined by estimates of the strength of the semantic relationships between the concepts within it. Originally, these estimates were obtained from human subjects, but more recent research has utilized estimates of the relatedness between terms that are derived automatically from large text corpora, using methods of distributional *semantics* (Schvaneveldt and Cohen 2010).

Distributional Semantics

A number of methods have been developed to automate the derivation of similarity metrics between terms based on distributional statistics of unannotated electronic text (for reviews, see (Cohen and Widdows 2009; Turney and Pantel 2010)). Spatial semantic models define terms as vectors in high dimensional space according to the distribution of their occurrence across a large text corpus. Semantic space models use different approaches to derive this multi-dimensional space, with an important fundamental difference having to do with the unit of text that is considered an independent context. For example, in Latent Semantic Analysis (LSA), each document in a text collection is considered as a unique context (Landauer and Dumais 1997), so a word is initially represented as a vector with a coordinate for each document in the collection, and values that are derived from counts of the number of times this word occurs in each document (the matrix is subsequently decomposed for the purpose of dimension reduction, which permits identification of second-order relationships between words that don't occur directly). In contrast, the Hyperspace Analogue to Language (HAL) (Lund and Burgess 1996), uses a sliding window around a term of interest as a context (as do some more recent approaches, such as neural word embeddings (Mikolov et al. 2013)), and counts the number of times terms co-occur with one another within this window as it moves through the corpus. So the initial representation of a word is a vector with a coordinate for each term in the vocabulary. In either case, the coordinates of a term vector in semantic space are determined by the distributional statistics for this term, such that similar vector representations are created for terms that occur in similar contexts. Evidence suggests that the semantic relatedness measured using LSA and other distributional models agrees with human estimates, and can be used to obtain human-like performance in a number of cognitive tasks (see for example (Lund and Burgess 1996; Landauer et al. 2006)). In the context of health-related online content, HAL has been combined with supervised machine learning algorithms, which learn to assign labels to vector representations of text from a training set that includes human annotation, to automatically classify consumer health webpages based on language use patterns (Chen et al. 2008). Estimates of distributional similarity derived from context vector representations of words have also been used to identify words within online discussions that fall into particular semantic categories (such as "medications" (Elhadad et al. 2014)), based on their similarity to an example seed term. More recently, other distributional models, specifically probabilistic topic models (which are part of a family of widely-used generative probabilistic models, that includes probabilistic Latent Semantic Analysis (Hofmann 2001) and Latent Dirichlet Allocation (Blei et al. 2003; Griffiths and Steyvers 2002)) and neural word embeddings (a neural-network based approach learns to predict term-to-context relationships, and has gained popularity in recent years) were evaluated for their utility as features for machine learning classifiers on the task of reproducing qualitative coding assigned to content mined from the cancer-related online forum, breastcancer.org (Zhang et al. 2016). The best performance on this task was obtained by using neural embeddings as features for a convolutional neural network (micro-averaged F-measure of 65.4). On the basis of these results, the qualitative analysis was extended to a larger number of messages using the trained classifier, permitting an innovative analysis of the evolution of topic trends over time. While methods for automated analysis of free text are still evolving, these methods have the capability to deal with large amounts of data generated by social media. Our own efforts in this area have also used distributional semantics and machine learning to extend the range of qualitative analysis. However, this work has focused on the integration of estimates of relatedness with quantitative network models, and will be discussed subsequently in the chapter once we have covered this methodological territory.

15.4.4 Quantitative Models of Network Science

Social network analysis has been widely used to examine network influence on individual behavior (Valente et al. 2004). For instance, friends' influence was operationalized as the extent to which adolescents are exposed to friends who use substances, and association with self-use was tested (Ali and Dwyer 2010; Crosnoe 2006; Crosnoe et al. 2004; Ennett et al. 2006; Fujimoto and Valente 2012a, b; Urberg et al. 1997).

Methodologically, the network exposure model (Burt 1987; Marsden and Friedkin 1993; TWTW 1995; Valente 2005) has been a workforce for modeling theories of social contagion (*i.e.*, behaviors change as a result of patterns of friendship relations) based on one-mode network. It specifies the appropriate weight matrix (W) for various network influence processes (Leenders 2002), and statistically testing network effects on individual behaviors. The general formula of network exposure E_i is defined as (TWTW 1995):

$$E_i = \frac{\sum_j W_{ij} Y_j}{\sum_j W_{ij}} \tag{15.1}$$

where W_{ij} is a social influence weight matrix, and Y_j is a vector of alter j's behavioral attribute ($i = 1, \ldots, N$; $j = 1, \ldots, N-1$, $i \neq j$). Exposure is calculated by matrix-multiplying the weight matrix, W (representing social influence matrix), by a vector indicating whether or not each alter j engages in behavior of interest (dichotomous variable coded as 0 or 1). The level of ego's exposure to behavior of interest is measured as the proportion of alters who engage in that behavior in an ego's network.

Affiliation Exposure Model

Network exposure model has been extended to the two-mode version, and the "affiliation exposure model (AEM)" was developed (Fujimoto et al. 2011, 2012), to model affiliation-based social contagion. AEM is designed to measure the degree to which individuals are exposed to behaviors of others through affiliating with the same settings/places. It uses pairwise relationships among individuals formed by sharing at least one settings/places. Mathematically, this requires the conversion of the original two-mode network data, **A**, into a one-mode projection of the *actor-by-actor* c-affiliation data, **C** (=**AA′**). In such converted one-mode co-membership network, each pair of actors is connected if they share at least one common place, thus representing the affiliation-based social influence.

By multiplying C_{ij} by each co-participant's attribute y_j and normalizing it by row-sum C_{i+} (ignoring the diagonal), the resulting affiliation exposure vector of $\underline{F}$ is defined as follow:

$$\underline{F} = \frac{\sum_{j=1} C_{ij} Y_j}{\sum_{j=1} C_{ij}} \qquad for \;\; i, j = 1, \cdots, N \quad i \neq j \tag{15.2}$$

AEM is a promising development since it does not require the collection of traditional network data (sociometrics), but instead just the collection of information relevant to attendance at specific places (*i.e.*, venues, and online meetings), something which is easily implemented in the survey format items common to public health research. Thus, AEM has the potential to contribute to enhancing the utility of network science in public health research areas that have not traditionally collected social network data. AEM has been used by other studies in different domains, including an online community for smoking cessation intervention (Myneni et al. 2015), the diffusion of the ratification of the WHO Framework Convention on Tobacco Control among countries (Wipfli et al. 2010), and a network among gangsters through criminal activities (Papachristos et al. 2015).

Decomposed Network Exposure Model

Different types of relationships may have different levels of influence with others in various health-related behaviors. Standard network exposure models handle a single type of relationship in studying the network influence on individuals' behaviors. To address this issue, some variants of network exposure models, called decomposed network exposure model (Fujimoto et al. 2013), have been introduced that are capable of handling multiple relationships by methodologically segregating the overlapped effect of one type and another types of networks, from non-overlapped effect, on individuals' behavior. For instance, one type of affiliation exposure may be overlapped with another type of network influence based on one-mode network or with another type of network influence based on two-mode network of different type. The decomposed network exposure model allows us to partition the model into two separate models.

$$\underline{D}_{(1)} = \frac{\sum_{j=1} X_{ij} C_{ij} Y_j}{\sum_{j=1} X_{ij} C_{ij}} \quad for\, i, j = 1, \cdots, N, \quad i \neq j \tag{15.3}$$

Mathematically, Eq. (15.3) was computed by element-wise product of the co-affiliation matrix C by an adjacency matrix X (representing one-mode network), and then row-normalized it and matrix-multiplied it by the behavioral vector of alters y_j. To compute Eq. (15.4), we subtracted an adjacency matrix X from a unit matrix with all elements being one, and everything else being identical to the computation of Eq. (15.3), which is defined in the following formula:

$$\underline{D}_{(2)} = \frac{\sum_{j=1} \left(1 - X_{ij}\right) C_{ij} Y_j}{\sum_{j=1} \left(1 - X_{ij}\right) C_{ij}} \quad for \quad i, j = 1, \cdots, N, \; i \neq j \tag{15.4}$$

Several empirical studies have applied this model to research on health behavior and public health by decomposing the effect of activity members who are also friends and activity members who are not friends, on adolescent substance use behavior (Fujimoto and Valente 2013).

To summarize, network exposure models enable us to measure a given form of network influence and test its effect on individual behavior to explain how new ideas and practices spread through social networks. One of the limitations in this model would be that network exposure model makes an independence assumption for the error term, which may not hold true for network data.

Network Autocorrelation Model

Network autocorrelation model (Doreian 1980, 1989; Doreian et al. 1984; Dow 1984; Ord 1975), also called the Network effects model (Doreian et al. 1984), assumes that endogenous network variables are correlated with the error term ε, and therefore the standard OLS regression yields biased and inconsistent estimates for both autocorrelation parameter and regression coefficient (Dow 2007; Johnston 1984).

However, these models are limited in the application to health behavioral and public health research, perhaps since all of these have been limited to modeling continuous-scale dependent variable. For instance, the outcome scale of substance use is usually measured by categorical scale (dichotomous or ordered categorical) since the frequency of a given substance use (such as alcohol use, cigarette smoking, marijuana use) is rarely normally distributed, and existing methods of network autocorrelation model does not handle categorical data analysis, except for a few network studies that uses two-stage, least-square regression with school level fixed effects using longitudinal data (Ali and Dwyer 2009, 2010).

Longitudinal Statistical Model

A series of Framingham network studies estimated social contagion effects by modeling the spread of obesity (Christakis and Fowler 2007), smoking (Christakis and Fowler 2008), alcohol use (Rosenquist et al. 2010), happiness (Fowler and Christakis 2008), depression (Rosenquist et al. 2011) and others, across network ties. These studies specified longitudinal regression models where the ego's (i.e. focal node) outcome status at any given time point $t + 1$ was a function of the ego's outcome status at the previous time point t, the alter's outcome status at times t and $t + 1$, controlling for various ego's attributes. Here, a significant coefficient for the alter's (i.e. nodes directly connected to focal node) outcome status at time $t + 1$ represents either an alter's outcome affected an ego's outcome (social contagion) or alter and ego experienced contemporaneous events affecting both outcome statues (environmental confounding) (Christakis and Fowler 2008). The model by Framingham longitudinal statistical model differs from network exposure models in that this model specifies dyadic tie as a unit of analysis, which may worsen the problem of model inconsistency especially for mutual ties. On the other hand, network exposure model specified individuals a unit of analysis (by measuring summary statistics of outcomes of alters who are connect to an individual, and which is computed for each individual).

However, the statistical problems inherent in using network data (or dyadic tie information) in regression models have been catalogued in response to these criticisms of social contagion, and a number of articles were published that discussed these critiques on the statistical procedures (Lyons 2011; Fowler and Christakis 2008; Christakis and Fowler 2013; Cohen-Cole and Fletcher 2008; Halliday et al. 2007; Shalizi 2012; VanderWeele et al. 2012), or proposed some remedies for potential problems of model inconsistency and estimation method used in modeling social contagion such as using instrumental variable (Halliday et al. 2007) or lagging of the alter's stage by an additional period (VanderWeele et al. 2012).

Exponential Random Graph Models (ERGMs)

Exponential Random Graph Models (ERGMs) (Frank and Strauss 1986; Robins et al. 2007; Wasserman and Pattison 1996; Hunter 2007; Wang et al. 2013) are capable of addressing this issue by treating the network itself as endogenous, and viewing the overall network structure as collections of local network processes represented by various structural configurations. ERGMs are designed to stochastically model the formation of network ties and test hypotheses about both local configurations (such as reciprocity, transitivity, etc.) and the distribution of nodal attributes (such as gender, drinking, smoking) within the network. Empirical network studies on health behavioral and public health research that applied ERGMs include assessing community-based participatory action designed to reduce cancer disparities (Valente et al. 2010), modeling peer selection mechanism based on adolescent's obesity or substance use behavior (Valente et al. 2009), and HIV risk transmission networks through venue affiliation among drug-using male sex workers (Fujimoto et al. 2015).

To conclude, the methods discussed so far provide valuable insights into the ways in which multidisciplinary techniques from behavior science, psychology, computer science, and network science can be used to conduct "social listening", thus enabling us to understand and model behavior change, social influence, information spread. Tools such as UCINET (Borgatti and Everett 2002), NetDraw (Borgatti 2002), Gephi (Bastian et al. 2009), and Cytoscape (Shannon et al. 2003) offer GUI based network modeling capabilities. These methodologies are promising in offering insights into designing and implementing network interventions that affect people's health behavior and public health problems, and translating them into practical public health interventions. In the next section of the chapter, we discuss our own efforts of applying the methods discussed so far to analysis of peer-to-peer communication in a health-related online social media platform.

15.5 Leveraging Social Media to Model and Change Behavior: A Case Study of Smoking Cessation

15.5.1 Introduction to QuitNet

QuitNet is one of the first online social networks for health behavior change, and has been in continuous existence for the past 17 years. It is widely used with over 100,000 new registrants per year (www.QuitNet.com). QuitNet has members who are current and former smokers seeking to quit or stay abstinent. The members are globally distributed and come from over 160 countries including Canada, the United Kingdom, Australia and South Africa. QuitNet's website incorporates the United States Public Health Service guidelines for best practice and includes diagnostic tools, social support from peers and experts, and pharmacotherapy (Cobb et al. 2005). It is available to smokers through two main channels: free public internet access and paid contracts. Both versions operate in the same environment and have a single support community, therefore, regardless of the means by which users access QuitNet, they all participate in the same online community. Research materials used for this study were extracted from the publicly available version.

Previous studies on QuitNet indicated that participation in the online community was strongly correlated with abstinence (Graham et al. 2007). Our studies outlined in this chapter include datasets drawn from a previously studied quality improvement database, and is comprised of de-identified messages in the public threaded forums, in which participants post messages and reply directly to each other. We have based our initial exploratory work on a de-identified 10 year data set of the original version of QuitNet, spanning 1996–2015 including and containing more than 400,000 individuals, 10 million inter-member communications and 194 million discrete behavioral observation points, including more than 500,000 that are specific to smoking behavior or medication use. All messages were stripped of identifiers but re-coded for ego id (the individual posting) and alter id (the individual whose message is being replied to), self-reported smoking status of sender and receiver ('0' for aspiring quitter, '1' for current smoker/non-quitter), date and position within the thread. QuitNet members were classified into four groups based on their self-reported smoking status. The classification criteria were as follows:

- 0: Members who were smokers throughout the study period (current smokers)
- 1: Members who stayed abstinent during the entire study period (ex-smokers)
- 0–1: Members who switched their status from smokers to ex-smokers (successful quitters)
- 1–0: Members who altered status from ex-smokers to smokers (relapsers)
- Other: Members who changed their smoking status multiple times (frequent relapsers)

In the next sections of the chapter, we describe a series of our prior studies conducted on QuitNet user communication using a variety of methods described in Sect. 15.4.

15.5.2 Qualitative Analysis of QuitNet Communication Content

In this section, we describe the results derived from a grounded theory-based (Strauss and Corbin 1998) content analysis of QuitNet messages. The findings derived from this analysis using our method are then interpreted in the light of existing behavior change theories, in an attempt to understand the interplay between the behavior changes facilitated by Web 2.0 based interventions and existing health behavior models. This analysis enhances our understanding of the applicability of behavior change theories (discussed in Sect. 15.4) which were formulated in the context of face-to-face communication using laboratory-based social science approaches, in the context of online social relationships.

Description of Methods

A grounded theory approach (Strauss and Corbin 1998) was used to analyze QuitNet data to understand the core concepts, the interrelations among concepts and the roles played by these concepts in an individual's smoking cessation activity. The first step in the coding process involved open coding, where a line-by-line analysis was performed on the messages to derive abstract concepts from the data. Each message was reviewed, noting pertinent smoking cessation related concepts in terms of general open codes which were generated dynamically as the data were reviewed. Examples of open codes included "statistics", "crave", "pregnancy", "boredom", "temper", "patch", and "pledge". This process was repeated until no new concepts were produced from the dataset. Appropriateness of code assignment was ascertained using constant comparison, where instances of codes were compared in an iterative manner to make sure they reflected the same concept. The second step was performed by re-organizing and re-grouping the open codes using axial coding. Axial coding allowed for the identification of unifying, repeated patterns underlying the concepts and their relationships, thereby revealing core themes relevant to smoking cessation. Examples of core themes include "Family and friends", "Obstacles", and "Traditions". Initial coding was performed manually, and later the NVivo software suite for qualitative analysis was used to analyze themes and their patterns of occurrence in the data. A total of 585 messages were analyzed, revealing 43 distinct concepts. Furthermore, the analysis was carried out for an additional 210 messages to ensure no new concepts emerged. This qualitative coding allowed for an in-depth evaluation of the interactions among people in the QuitNet social platform and thereby a deeper understanding of the behavior change processes that QuitNet users undergo when attempting to cease smoking. Further, these themes were mapped to theoretical constructs and taxonomy techniques outlined in behavior change theories discussed in Sect. 15.4.1.

Summary of Results and Conclusions

Communication themes ranged from discussions on nicotine replacement therapies to stress management strategies. QuitNet users posted messages seeking help to fight a craving or confessing to a relapse. Importantly, the analysis revealed aspects of community-specific culture such as "Saturday night bonfires"—where unsmoked cigarettes are thrown into a digital fireplace during a virtual gathering and "early morning weather updates" when the users reaffirm their willingness to not smoke.

Fig. 15.1 QuitNet communication themes

Messages also indicated issues with building trust between members. In the case of QuitNet, activities such as pledges and bonfires emerged from within the community and each of those events marks a specific aspect of the smoking-cessation process. In addition to emphasizing progress and positive aspects of smoking cessation, focus on community-building and social togetherness (e.g., bonfires) have helped members adhere to their quit attempts. Like any other virtual community, most content embeds aspects of social support. In addition to support, several other sociobehavioral elements related to behavior change theories were found in QuitNet messages. Our analysis revealed that most QuitNet themes (1) relate to important behavior change constructs belonging to multiple theories and (2) operationalize several techniques outlined in the behavior change taxonomy, thus highlighting the need for empirically-grounded behavior change interventions. Figure 15.1 presents the prevalence of various content types in QuitNet user communication. Detailed description of the thematic definitions, their mapping to the theories and taxonomy techniques can be found in (Myneni et al., 2012, 2013, 2016a).

Qualitative methods form a very important toolkit to conduct nuanced analysis of health-related communications in online platforms. Use of grounded theory analysis has allowed us to develop thematic representations of QuitNet messages that are empirically driven and not theoretically biased. Subsequently, comparison analysis consisting of (1) sociobehavioral constructs from existing behavior change theories and (2) theoretically linked taxonomy of behavior change techniques allowed us to understand the theoretical roots and operational features of consumer-driven QuitNet communication. The methodological process itself is informative, comprehensive, and generalizable, while being both empirically grounded and theoretically aligned.

Future Directions

As part of our ongoing studies, we have extended our empirical and theory-driven qualitative analysis to model user communication in social media platforms specifically designed for chronic diseases and healthy living. Such cross-community and cross-behavior analysis will allow us to generalize the observed phenomena and understand relationship between communication themes, socio-behavioral theories, and user engagement attributes underlying behavior change and chronic disease management.

In the next section of the paper, we describe how we applied automated text analysis methods from distributional semantics in conjunction with machine learning algorithms to enable high-throughput analysis of online social media communications. Such methods facilitate resource-optimized extension of qualitative analysis to large-scale digital health data. This in turn can extend the research and application frontiers of social media, thereby further enhancing their positive impact on health-related behaviors.

15.5.3 Automated Text Analysis of QuitNet

As web forums are the predominant modes of communication in social media communities, recent advances in automated text analysis allow for large-scale analysis of the content of peer-to-peer interactions. Semantic space models, methods of distributional semantics in which both terms and larger units of text are represented in a high-dimensional vector space, have been applied to peer-to-peer interactions in online communities (McArthur et al. 2006; Mc Arthur and Bruza 2002, 2003). The methods of automated text analysis we have employed infer measures of the relatedness between passages of text from the distributional statistics of terms in a large text corpus.

Description of Methods

In our prior work we drew on external distributional information, from the Touchstone Applied Science Associated (TASA) corpus (Landauer et al. 1997), a collection of 37,657 articles designed to approximate the average reading of an American college freshman, to account for terse semantic context available in social media postings. We then used LSA (Landauer et al. 1998) to derive vector representations of terms in the TASA corpus, such that terms with similar distributions would have similar vector representations, and measured similarity between vectors using the cosine metric. In addition, Reflective Random indexing (RRI) a variant of Random Indexing (Kanerva et al. 2000) which was developed to recognize meaningful relationships between terms without requiring they co-occur directly (Cohen et al. 2010). LSA and RRI were performed using the Semantic Vectors package (Widdows and Ferraro 2008; Widdows and Cohen 2010), an open source package for distributional semantics. The log-entropy weighting metric was used, and terms occurring on the stopword list distributed with the General Text Parser software package (Giles et al. 2003) were ignored. This stopword list consists of frequently occurring terms that carry little semantic content. Subsequently, representations of the messages in the QuitNet corpus were generated using an iterative approach (Vasuki and Cohen 2009). In order to use these generated vectors to support automated coding of QuitNet messages, we conducted a series of studies using (a) key word based modeling, (b) nearest neighbors approach, (c) machine learning algorithms using Weka, a popular open source machine learning toolkit (Hall et al. 2009). Detailed explanation of the methods can be found in (Myneni et al. 2012, 2015, 2016b). Accuracy and reliability metrics were employed to assess the performance of the automated methods (Myneni et al. 2012, 2015).

Summary of Results and Conclusions

(a) Incorporation of external corpus: LSA in conjunction with a range of classifiers for categorization of the QuitNet set has shown significant improvements with TASA incorporation. Accuracy measures using Key word models (theme-specific manual inspection, see (Myneni et al. 2012)) and nearest neighbors approach (leave one out cross validation, see (Myneni et al. 2016b)) without TASA pre-training are 0.48 and 0.53 respectively. With TASA pre-training the accuracy improved to 0.64 and 0.74 for key words model and nearest neighbors approach respectively (Myneni et al. 2016b).

(b) Accuracy and Reliability: The optimal F-measures, precision and recall metrics for the cross validation were achieved with the application of reflective random indexing in conjunction with J48 decision tree built within Weka: Recall = 0.76, Precision = 0.78, and F-measure = 0.77. In addition, reliability measurements between two human coders and machine coding were as follows. The reliability between coder 1 and the system is 0.71, coder 2 and the system is 0.74, and these values indicate average coder-system reliability approached coder-coder agreement of 0.74 (Sridharan et al. 2016).

Future Directions

Communication exchanges in online communities are time-stamped which facilitates tracking of semantic changes in the messages exchanged by a user pair or a group of users over a time period. Utilization of time series analysis (O'Connor et al. 2010) and novel methods of encoding sequence within distributional semantic models such as (Widdows and Cohen 2016) may enable longitudinal semantic modeling of QuitNet communication events, thus offering deeper insights into the evolution of behavior change processes as QuitNet users attempt to stay abstinent from smoking.

15.5.4 Network Models of QuitNet

To date, most network analysis studies on health-related online social networks have focused primarily on exploring the structural and functional composition of networks without considering communication content. Efforts have been made to evaluate the quality (in terms of semantic features such as syntactic and semantic complexity, punctuation, and grammaticality) of content in social media (Agichtein et al. 2008), facilitate social tagging, where users to annotate, categorize and share their web content using short textual labels (Fu and Kannampallil 2009). Few efforts have been made to bridge content-rich and content-free analyses to characterize communication in social networks. In the following section, we describe ways in which content-inclusive network analysis has been conducted in our own work on QuitNet.

Description of Methods

Network models were created to characterize content-specific topology of QuitNet user communication. QuitNet users were represented as nodes and communication attributes (frequency, semantic similarity) were used to represent edges within the network. A variety of metrics (e.g. degree, modularity, clustering, pathlength) (Valente 2010) to identify communication-specific factors underlying QuitNet users. We used available open-source network simulation software, Gephi (Bastian et al. 2009) to model these content-specific networks. Further, we have framed communication themes as events and network members as actors, forming two-mode affiliation network data. Subsequently, the two-mode data derived were as illustrated in Table 15.1. The information in the table was based upon the categorization of messages obtained in Sects. 15.5.1 and 15.5.2. Thus, network members and their themes formed the two modes for network analysis using affiliate exposure models. Given that co-participation in our study is dependent on content of communication, this allows for the characterization of the role of content-specific social influence patterns underlying peer-to-peer communication. Consequently, we evaluated the extent to which "membership" in specific content type is predictive of smoking cessation and effective diabetes management.

Summary of Results and Conclusions

This work yielded insights that are interpretable and actionable, enabling the identification of content-specific opinion leaders (high degree nodes indicated by their large size) and their distribution (same color nodes implies clustering into a sub community) as shown in Figs. 15.2 and 15.3. Using coaffiliation networks (shown

Table 15.1 Illustration of two-mode data (Row mode represents QuitNet users and Column mode refers to the QuitNet communication themes)

QuitNet member	QuitNet theme	Social support	Obstacles	Rewards
ID00000XX		1	1	0
ID1111XX1		1	1	0
ID112233X		0	1	1
IDXXX2221		1	0	1

Fig. 15.2 Content-specific QuitNet networks

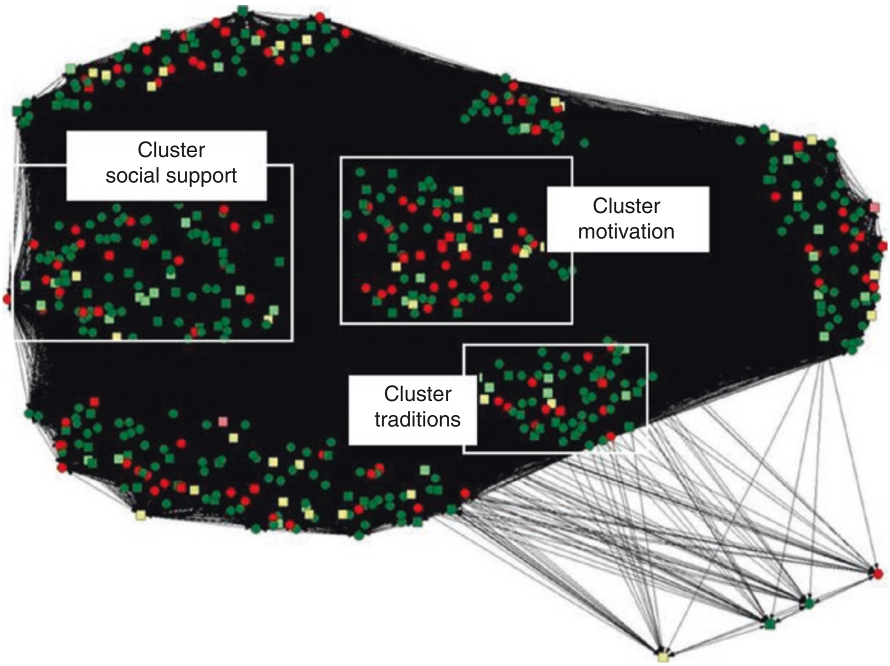

Fig. 15.3 QuitNet coaffiliation network

in Fig. 15.3) we examined the clusters among users based on joint affiliation with at least one common theme. Results revealed similar communication content exchange by users belonging to a common cluster. A two-mode version of network autocorrelation model developed was used to estimate the effect of content-specific affiliation exposure on individuals' abstinence status. The autocorrelation parameter estimate showed that as QuitNet users were more exposed to other users who stay abstinent and share community-centric themes (Social support, Traditions, Motivation), they were more likely to stay abstinent themselves ($b = 0.041$, $P < 0.01$) (Myneni et al. 2015), to the extent that self-reports are accurate.

Future Directions

Current network models attempt to understand behavioral diffusion by analyzing frequency of communication without considering its content. Our prior work on content-inclusive network analysis focused on cross-sectional integration of content with network structure and did not account for network dependencies beyond dyad-level (Myneni et al. 2015) (i.e. who communicates with whom). This is a significant limitation and our ongoing work attempts to negotiate this by modeling networks using ERGMs discussed in Sect. 15.4. Such methods will allow us to incorporate user-theme and user-user relationships to model content-specific social influence patterns. In addition, longitudinal network models can be used to examine content-specific topologies and temporal trends.

15.5.5 Implication for Public Health Interventions

Content-inclusive network analysis of social media communication as facilitated by methods discussed in the chapter so far provides rich empirical evidence that form the basis for data-driven health promotion ventures that provide new directions for research on interventions that health researchers and technology developers can undertake to change human health behavior. Application areas of targeted network interventions include (a) identification of opinion leaders, clusters, and group-specific opinion leaders, (b) "rewiring" networks to improve or reduce network cohesion, and (c) network-attribute interventions (Valente 2012). Content attributes can be used to derive network-attribute interventions, where members exchanging messages related to a particular theme will be segmented as a group to harness the positive effects of their social influence. Examples of these approaches include identifying content-specific key players and creating mentor-mentee relationships based on the needs of the mentee and interests of the mentor. The one-mode network structures obtained from formal network analysis using Gephi reveal differences across themes, the most striking of which is the difference in the high-degree nodes across the themes, which indicate those users with the most connections with whom they discuss content related to a particular theme. Consequently, these high-degree nodes represent the opinion leaders of the network with respect to those particular themes. Identifying key players within groups was shown to be one important step for effective in tobacco control (Puska et al. 1986). Opinion leaders play a pertinent role in social mobilizations and social networks, they act as gatekeepers for interventions, help change social norms, and accelerate behavior change (Obregón et al. 2009; Valente and Pumpuang 2007). The opinion leaders identified through our work discussed in Sect. 15.5.3 were within a group of members exchanging information related to a specific topic of interest such as "Social Support", "Traditions", and "Progress". For example, the identification of those network members who are key players in providing "Relapse assistance" and "Motivation" can help us make the right connections with users discussing about "Craves", thus improving the network's assistance to its members. This new knowledge about content-specific opinion leaders can be transformed into a content-sensitive targeted intervention by incorporating new support features into a social network for providing guidance information to network users with respect to content variety and content-specific opinion leaders. For example, if a network member exchanged messages related to "Progress", then that member can be directed toward similar content types and the opinion leaders for that particular content. In addition, the "Progress"-related opinion leaders can be alerted about the new member to facilitate a connection between this member and an opinion leader. Similarly, if a member posts messages that indicate "Conflict", trust-related issues with another member, then directing them toward messages indicating "Social support" and "Motivation" may be of assistance.

In terms of factors affecting content-based network influence, results indicate that exposure to abstinent members exchanging content related to group-centric inter-personal themes (e.g. "Social support", "Traditions") tend to stay abstinent from smoking behavior. Therefore, online interventions can incorporate an explicit display of member profiles contributing to such content to enhance affiliations to these people. In the context of offline interventions at population level, public support messages incorporating content features highlighting the need to seek social support and be part of a group-based smoking cessation endeavor can help the general public (confronting similar issues) become involved in a support community to sustain abstinence from smoking. Preliminary efforts involving the design of an empirically-informed social support platform inspired by social media analytics can be found here (Myneni and Iyengar 2016).

In summary, online social networks have been gaining in popularity and present health researchers with a unique opportunity to understand human behavior change and deliver scalable and sustainable interventions. However, as demonstrated by QuitNet, these venues can also provide a forum for a community of dedicated users to assist one another in the pursuit of better health, an activity that ultimately has societal benefit beyond the users of QuitNet itself. The development of better tools to analyze social network content of this nature allows us a greater understanding of the ways in which such social networks mediate behavior change, thereby providing us with the opportunity for empirically-grounded interventions to further assist these communities with the attainment of their laudable goals. Content-based network analysis of QuitNet, made feasible by large-scale qualitative analysis using automated methods, has been shown to yield content-specific tailoring strategies that can be used for health promotion and behavior change.

Discussion Questions

1. Text mining and network analysis are important tools for social media analysis. How and why do these methods inform our understanding of human behavior and intervention design for behavior change?
2. Consider a public health behavior change challenge. Identify social media data sources for the behavior under question. Discuss user security and privacy intricacies associated with applying the social media analytics discussed in the chapter.

Acknowledgements Research reported in this publication was supported in part by the National Library of Medicine of the National Institutes of Health under Award Number 1R21LM012271-01, National Library of Medicine Grant Number 1R01LM011563, and UTHealth Innovation for Cancer Prevention Research Pre-doctoral Fellowship, The University of Texas School of Public Health-Cancer Prevention and Research Institute of Texas grant RP101503. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health and the Cancer Prevention and Research Institute of Texas. We would like to express our sincere gratitude to our collaborator Dr. Nathan K. Cobb for providing us with de-identified data from QuitNet platform. We would like to thank Tom Landauer for providing us with the TASA corpus, and contributors to the Semantic Vectors open source package, in particular Adrian Kuhn and David Erni, the contributors of the sparse SVD implementation we used for the LSA package.

References

Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol. 2008;27(3):379.

Agichtein E, Castillo C, Donato D, Gionis A, Mishne G. Finding high quality content in social media. Paper presented at the Proceedings of the international conference on Web search and web data mining, 2008.

Ahmed MY, Kenkeremath S, Stankovic J. Socialsense: a collaborative mobile platform for speaker and mood identification. In: European Conference on Wireless Sensor Networks [Internet]. Springer; 2015 [cited 2016 Jun 24]. p. 68–83. http://link.springer.com/chapter/10.1007/978-3-319-15582-1_5

Ajzen I. From intentions to actions: a theory of planned behavior. In: Action control [Internet]. Springer; 1985 [cited 2016 Jun 24]. p. 11–39. http://link.springer.com/chapter/10.1007/978-3-642-69746-3_2

Albarracin D, Johnson BT, Fishbein M, Muellerleile PA. Theories of reasoned action and planned behavior as models of condom use: a meta-analysis. Psychol Bull. 2001;127(1):142.

Alexander C, Piazza M, Mekos D, Valente T. Peers, schools, and adolescent cigarette smoking. J Adolesc Health. 2001;29(1):22–30.

Ali MM, Dwyer DS. Estimating peer effects in adolescent smoking behavior: a longitudinal analysis. J Adolesc Health. 2009;45(4):402–8.

Ali MM, Dwyer DS. Social network effects in alcohol consumption among adolescents. Addict Behav. 2010;35(4):337–42.

Aral S, Muchnik L, Sundararajan A. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. Proc Natl Acad Sci. 2009;106(51):21544–9.

Bachmann A. Towards smartphone-based sensing of social interaction for ambulatory assessment. In: Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on

Wearable Computers [Internet]. ACM; 2015 [cited 2016 Jun 24]. p. 423–428. http://dl.acm.org/citation.cfm?id=2801642
Bandura A. Social foundations of thought and action: A social cognitive theory [Internet]. Prentice-Hall, Inc.; 1986 [cited 2016 Jun 24]. http://psycnet.apa.org/psycinfo/1985-98423-000/
Bandura A. Social cognitive theory of mass communication. Media Psychol. 2001;3(3):265–99.
Bandura A. Social cognitive theory. In van Lange PAM, Kruglanski AW, Higgins ET, editors. Handbook of social psychological theories. London: Sage; 2011. p. 349–73.
Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. ICWSM. 2009;8:361–2.
Blei DM, Ng AY, Jordan MI. Latent Dirichlet allocation. J Mach Learn Res. 2003;3:993–1022.
Borgatti SP. NetDraw: graph visualization software. Harvard: Analytic Technologies; 2002
Borgatti SP, Everett MG, Freeman LC. Ucinet for Windows: software for social network analysis. 2002 [cited 2016 Jun 25]; http://www.citeulike.org/group/11708/article/6031268
Burt RS. Social contagion and innovation: cohesion versus structural equivalence. Am J Sociol. 1987;92:1287–335.
Cavallo DN, Tate DF, Ries AV, Brown JD, DeVellis RF, Ammerman AS. A social media-based physical activity intervention: a randomized controlled trial. Am J Prev Med. 2012;43(5):527–32.
Cavallo DN, Chou W-YS, McQueen A, Ramirez A, Riley WT. Cancer prevention and control interventions using social media: user-generated approaches. Cancer Epidemiol Biomark Prev. 2014;23(9):1953–6.
Centola D. The spread of behavior in an online social network experiment. Science. 2010;329(5996):1194–7.
Centola D. Social media and the science of health behavior. Circulation. 2013;127(21):2135–44.
Chen G, Warren J, Evans J. Automatically generated consumer health metadata using semantic spaces. In Wollongong, NSW, Australia: Australian Computer Society, Inc.; 2008 [cited 2008 Dec 15]. p. 9–15. http://portal.acm.org/citation.cfm?id=1385089.1385093
Chou W-YS, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. J Med Internet Res. 2009;11(4):e48.
Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. N Engl J Med. 2007;357(4):370–9.
Christakis NA, Fowler JH. The collective dynamics of smoking in a large social network. N Engl J Med. 2008;358(21):2249–58.
Christakis NA, Fowler JH. Social contagion theory: examining dynamic social networks and human behavior. Stat Med. 2013;32(4):556–77.
Chuang KY, Yang CC. A study of informational support exchanges in MedHelp alcoholism community. In: International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction [Internet]. Springer; 2012 [cited 2016 Jun 24]. p. 9–17. http://link.springer.com/chapter/10.1007/978-3-642-29047-3_2
Cobb NK, Graham AL, Bock BC, Papandonatos G, Abrams DB. Initial evaluation of a real-world Internet smoking cessation system. Nicotine Tob Res. 2005;7(2):207–16.
Cohen T, Widdows D. Empirical distributional semantics: methods and biomedical applications. J Biomed Inform. 2009;42(2):390–405.
Cohen T, Schvaneveldt R, Widdows D. Reflective Random Indexing and indirect inference: a scalable method for discovery of implicit connections. J Biomed Inform. 2010;43(2):240–56.
Cohen-Cole E, Fletcher JM. Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic. J Health Econ. 2008;27(5):1382–7.
Cohn AM, Hunter-Reel D, Hagman BT, Mitchell J. Promoting behavior change from alcohol use through mobile technology: the future of ecological momentary assessment. Alcohol Clin Exp Res. 2011;35(12):2209–15.
Coulson NS, Buchanan H, Aubeeluck A. Social support in cyberspace: a content analysis of communication within a Huntington's disease online support group. Patient Educ Couns. 2007;68(2):173–8.
Crosnoe R. The connection between academic failure and adolescent drinking in secondary school. Sociol Educ. 2006;79(1):44–60.

Crosnoe R, Muller C, Frank K. Peer context and the consequences of adolescent drinking. Soc Probl. 2004;51(2):288–304.

Doing-Harris KM, Zeng-Treitler Q. Computer-assisted update of a consumer health vocabulary through mining of social network data. J Med Internet Res [Internet]. 2011 [cited 2014 Mar 24];13(2). http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3221384/

Donovan S. The effectiveness of an internet support forum for carers of people with dementia: a pre-post cohort study. Prim Health Care. 2014;24(9):18.

Doreian P. Linear models with spatially distributed data spatial disturbances or spatial effects? Sociol Methods Res. 1980;9(1):29–60.

Doreian P. Network autocorrelation models: problems and prospects. In: Spatial statistics: past present future. 1989; p. 369–389.

Doreian P, Teuter K, Wang C-H. Network autocorrelation models some Monte Carlo results. Sociol Methods Res. 1984;13(2):155–200.

Dow MM. A biparametric approach to network autocorrelation Galton's problem. Sociol Methods Res. 1984;13(2):201–17.

Dow MM. Galton's problem as multiple network autocorrelation effects cultural trait transmission and ecological constraint. Cross-Cult Res. 2007;41(4):336–63.

Duggan M, Ellison NB, Lampe C, Lenhart A, Madden M. Social media update 2014. Pew Res Cent [Internet]. 2015 [cited 2016 Jun 24];9. http://www.foothillspresbytery.org/wp-content/uploads/sites/175/2015/07/Social-Media-Site-Usage-2014-_-Pew-Research-Centers-Internet-American-Life-Project.pdf

Elder JP, McGraw SA, Abrams DB, Ferreira A, Lasater TM, Longpre H, et al. Organizational and community approaches to community-wide prevention of heart disease: the first two years of the Pawtucket Heart Health Program. Prev Med. 1986;15(2):107–17.

Elhadad N, Zhang S, Driscoll P, Brody S. Characterizing the sublanguage of online breast cancer forums for medications, symptoms, and emotions. In: Proceedings AMIA Annual Fall Symposium [Internet]. 2014 [cited 2016 Jun 24]. https://pdfs.semanticscholar.org/1b03/f9b226a6778b42e951207f24b1bab59989b8.pdf

Ennett ST, Bauman KE, Hussong A, Faris R, Foshee VA, Cai L, et al. The peer context of adolescent substance use: findings from social network analysis. J Res Adolesc. 2006;16(2):159–86.

Facebook [Internet]. [cited 2016] Jun 24. https://www.facebook.com/facebook/info?tab=page_info

Fernandes R, Hatfield J, Job RS. A systematic investigation of the differential predictors for speeding, drink-driving, driving while fatigued, and not wearing a seat belt, among young drivers. Transp Res Part F Traffic Psychol Behav. 2010;13(3):179–96.

Fishbein M. A theory of reasoned action: some applications and implications. 1979 [cited 2016 Jun 24]; http://psycnet.apa.org/psycinfo/1982-21121-001

Fisher J, Clayton M. Who gives a tweet: assessing patients' interest in the use of social media for health care. Worldviews Evid-Based Nurs. 2012;9(2):100–8.

Fisher WA, Fisher JD, Rye BJ. Understanding and promoting AIDS-preventive behavior: insights from the theory of reasoned action. Health Psychol. 1995;14(3):255.

Fowler JH, Christakis NA. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. BMJ. 2008;337(2):a2338.

Frank O, Strauss D. Markov graphs. J Am Stat Assoc. 1986;81(395):832–42.

Freeman LC. Centrality in social networks conceptual clarification. Soc Networks. 1978;1(3):215–39.

Fu WT, Kannampallil TG, Kang R. A semantic imitation model of social tag choices. Paper presented at the Computational Science and Engineering, 2009.

Fujimoto K, Valente TW. Decomposing the components of friendship and friends' influence on adolescent drinking and smoking. J Adolesc Health. 2012a;51(2):136–43.

Fujimoto K, Valente TW. Social network influences on adolescent substance use: disentangling structural equivalence from cohesion. Soc Sci Med. 2012b;74(12):1952–60.

Fujimoto K, Valente TW. Alcohol peer influence of participating in organized school activities: a network approach. Health Psychol. 2013;32(10):1084.

Fujimoto K, Chou C-P, Valente TW. The network autocorrelation model using two-mode data: affiliation exposure and potential bias in the autocorrelation parameter. Soc Netw. 2011; 33(3):231–43.
Fujimoto K, Unger JB, Valente TW. A network method of measuring affiliation-based peer influence: assessing the influences of teammates' smoking on adolescent smoking. Child Dev. 2012;83(2):442–51.
Fujimoto K, Wang P, Valente TW. The decomposed affiliation exposure model: a network approach to segregating peer influences from crowds and organized sports. Netw Sci Camb Univ Press [Internet]. 2013 [cited 2014 Mar 24];1(2). http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3859688/
Fujimoto K, Wang P, Ross MW, Williams ML. Venue-mediated weak ties in multiplex HIV transmission risk networks among drug-using male sex workers and associates. J Inf [Internet]. 2015 [cited 2016 Jun 25];105(6). https://ajph.aphapublications.org/doi/full/10.2105/AJPH.2014.302474
Giles JT, Wo L, Berry MW. GTP (General Text Parser) software for text mining. Statistical data mining and knowledge discovery. 2003. p. 455–71.
Graham AL, Cobb NK, Raymond L, Sill S, Young J. Effectiveness of an internet-based worksite smoking cessation intervention at 12 months. J Occup Environ Med Am Coll Occup Environ Med. 2007;49(8):821–8.
Griffiths T, Steyvers M. A probabilistic approach to semantic representation. In: Proceedings of the 24th Annual Conference if Cognitive Science Society; 2002. p. 381–6.
Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA data mining software: an update. SIGKDD Explor Newsl. 2009;11(1):10–8.
Halliday TJ, Kwak S, et al.. Identifying endogenous peer effects in the spread of obesity. Unpublished Working Paper. University of Hawaii-Minoa [Internet]. 2007 [cited 2016 Jun 24]; http://core.ac.uk/download/pdf/7163214.pdf
Harper FM, Moy D, Konstan JA. Facts or friends? Distinguishing informational and conversational questions in social Q&A sites. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems [Internet]. ACM; 2009 [cited 2014 Mar 24]. p. 759–68. http://dl.acm.org/citation.cfm?id=1518819
Hawn C. Take two aspirin and tweet me in the morning: how Twitter, Facebook, and other social media are reshaping health care. Health Aff (Millwood). 2009;28(2):361–8.
Heaney CA, Israel BA. Social networks and social support. Health Behav Health Educ Theory Res Pract. 2008;4:189–210.
Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol. 2010;15(1):1–39.
Hochbaum G, Rosenstock I, Kegels S. Health belief model. U S Public Health Serv [Internet]. 1952 [cited 2016 Jun 24]. http://www.infosihat.gov.my/infosihat/artikelHP/bahanrujukan/HE_DAN_TEORI/DOC/Health%20Belief%20Model.doc
Hofmann T. Unsupervised learning by probabilistic latent semantic analysis. Mach Learn. 2001;42:177–96.
House JS, et al. Work stress and social support [Internet]. Addison-Wesley Pub. Co.; 1981 [cited 2016 Jun 24]. http://agris.fao.org/agris-search/search.do?recordID=US201300364231
House JS, Kahn RL, McLeod JD, Williams D. Measures and concepts of social support. 1985 [cited 2016 Jun 24]; http://psycnet.apa.org/psycinfo/1985-97489-005
Hunter DR. Curved exponential family models for social networks. Soc Netw. 2007;29(2): 216–30.
Hwang KO, Ottenbacher AJ, Green AP, Cannon-Diehl MR, Richardson O, Bernstam EV, et al. Social support in an Internet weight loss community. Int J Med Inform. 2010;79(1):5–13.
Johnston J, JE DN. Econometric methods. New York: McGraw Hill; 1984.
Kanerva P, Kristofersson J, Holst A. Random indexing of text samples for latent semantic analysis. In: Proceedings of the 22nd Annual Conference of the Cognitive Science Society; 2000. p. 1036.
Kaplan AM, Haenlein M. Users of the world, unite! The challenges and opportunities of Social Media. Bus Horiz. 2010;53(1):59–68.

Kietzmann JH, Hermkens K, McCarthy IP, Silvestre BS. Social media? Get serious! Understanding the functional building blocks of social media. Bus Horiz. 2011;54(3):241–51.

Killen JD, Robinson TN, Telch MJ, Saylor KE, Maron DJ, Rich T, et al. The Stanford Adolescent Heart Health Program. Health Educ Q [Internet]. 1989 [cited 2016 Jun 24]; http://psycnet.apa.org/psycinfo/1996-73085-001

Klemm P, Bunnell D, Cullen M, Soneji R, Gibbons P, Holecek A. Online cancer support groups: a review of the research literature. Comput Inform Nurs. 2003;21(3):136–42.

Korda H, Itani Z. Harnessing social media for health promotion and behavior change. Health Promot Pract. 2013;14(1):15–23.

Landauer TK. Handbook of latent semantic analysis. Mahwah, NJ: Lawrence Erlbaum Associates; 2006.

Landauer TK, Dumais ST. A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychol Rev. 1997;104:211–40.

Landauer TK, Laham D, Rehder B, Schreiner ME. How well can passage meaning be derived without using word order? A comparison of latent semantic analysis and humans. Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society; 1997 Aug 7–10; Stanford University. 1997

Landauer TK, Foltz PW, Laham D. An introduction to latent semantic analysis. Discourse Process. 1998;25:259–84.

Leenders RTA. Modeling social influence through network autocorrelation: constructing the weight matrix. Soc Netw. 2002;24(1):21–47.

Lund K, Burgess C. Producing high-dimensional semantic spaces from lexical co-occurrence. Behav Res Methods Instrum Comput. 1996;28:203–8.

Lyons R. The spread of evidence-poor medicine via flawed social-network analysis. Stat Polit Policy [Internet]. 2011 [cited 2016 Jun 24];2(1). http://www.degruyter.com/view/j/spp.2011.2.issue-1/spp.2011.2.1.1024/spp.2011.2.1.1024.xml

Marsden PV, Friedkin NE. Network studies of social influence. Sociol Methods Res. 1993;22(1):127–51.

Matsuo Y, Mori J, Hamasaki M, Nishimura T, Takeda H, Hasida K, et al. POLYPHONET: an advanced social network extraction system from the web. Web Semant Sci Serv Agents World Wide Web. 2007;5(4):262–78.

Mc Arthur R, Bruza P. 5. Dimensional representations of knowledge in online community. 2002. Doi: 10.1007/978-3-662-06230-2_8

Mc Arthur R, Bruza P. 6. Discovery of tacit knowledge and topical ebbs and flows within the utterances of online community. (2003). DOI: 10.1007/978-3-662-06230-2_9

McArthur R, Bruza P, Warren J, Kralik D. Projecting computational sense of self: a study of transition in a chronic illness online community. 2006. p. 91c.

McGloin AF, Eslami S. Digital and social media opportunities for dietary behaviour change. Proc Nutr Soc. 2015;74(2):139–48.

Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med. 2013;46(1):81–95.

Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. Proceedings of the First International Conference on Learning Representations (ICLR) 2013. Prepr ArXiv13013781 [Internet]. 2013 [cited 2016 Jun 24]; http://arxiv.org/abs/1301.3781

Myneni S, Iyengar S. Socially influencing technologies for health promotion: translating social media analytics into consumer-facing health solutions. In: IEEE Hawaii International Conference on System Sciences, 2016.

Myneni S, Cobb NK, Cohen T. Finding meaning in social media: content-based social network analysis of QuitNet to identify new opportunities for health promotion. Stud Health Technol Inform. 2012;192:807–11.

Myneni S, Iyengar S, Cobb NK, Cohen T. Identifying persuasive qualities of decentralized peer-to-peer online social networks in public health. In: Persuasive Technology [Internet]. Springer; 2013 [cited 2014 Mar 24]. p. 155–160. http://link.springer.com/chapter/10.1007/978-3-642-37157-8_19

Myneni S, Fujimoto K, Cobb N, Cohen T. Content-driven analysis of an online community for smoking cessation: integration of qualitative techniques, automated text analysis, and affiliation networks. Am J Public Health. 2015;105(6):1206–12.

Myneni S, Cobb N, Cohen T. In pursuit of theoretical ground in behavior change support systems: analysis of peer-to-peer communication in a health-related online community. J Med Internet Res [Internet]. 2016a [cited 2016 Jun 24];18(2). http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4756252/

Myneni S, Cobb N, Cohen T. Content-specific network analysis of peer-to-peer communication in an online community for smoking cessation. Paper to be presented at AMIA Annual Symposium Proceedings 2016b, Chicago, IL.

Nimrod G. Seniors' online communities: a quantitative content analysis. Gerontologist. 2010;50(3):382–92.

O'Connor B, Balasubramanyan R, Routledge BR, Smith NA. From tweets to polls: linking text sentiment to public opinion time series. ICWSM. 2010;11(122–129):1–2.

Obregón R, Chitnis K, Morry C, Feek W, Bates J, Galway M, et al. Achieving polio eradication: a review of health communication evidence and lessons learned in India and Pakistan. Bull World Health Organ. 2009;87(8):624–30.

Ord K. Estimation methods for models of spatial interaction. J Am Stat Assoc. 1975;70(349):120–6.

Papachristos AV, Wildeman C, Roberto E. Tragic, but not random: the social contagion of nonfatal gunshot injuries. Soc Sci Med. 2015;125:139–50.

Live better, together! | PatientsLikeMe [Internet]. [cited 2016] Jun 24. https://www.patientslikeme.com/

Perry CL, Kelder SH, Murray DM, Klepp K-I. Communitywide smoking prevention: long-term outcomes of the Minnesota Heart Health Program and the Class of 1989 Study. Am J Public Health. 1992;82(9):1210–6.

Poirier J, Cobb NK. Social influence as a driver of engagement in a web-based health intervention. J Med Internet Res. 2012;14(1):e36.

Prochaska JJ, Prochaska JO. A review of multiple health behavior change interventions for primary prevention. Am J Lifestyle Med. 2011;5(3):208–21.

Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot. 1997;12(1):38–48.

Puska P, Koskela K, McAlister A, Mäyränen H, Smolander A, Moisio S, et al. Use of lay opinion leaders to promote diffusion of health innovations in a community programme: lessons learned from the North Karelia project. Bull World Health Organ. 1986;64(3):437–46.

QuitNet [Internet]. [cited 2016] Jun 24. https://quitnet.meyouhealth.com/

Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? Transl Behav Med. 2011;1(1):53–71.

Robins G, Snijders T, Wang P, Handcock M, Pattison P. Recent developments in exponential random graph (p*) models for social networks. Soc Netw. 2007;29(2):192–215.

Rosenquist JN, Murabito J, Fowler JH, Christakis NA. The spread of alcohol consumption behavior in a large social network. Ann Intern Med. 2010;152(7):426–33. W141

Rosenquist JN, Fowler JH, Christakis NA. Social network determinants of depression. Mol Psychiatry. 2011;16(3):273–81.

Royce JM, Hymowitz N, Corbett K, Hartwell TD, Orlandi MA. Smoking cessation factors among African Americans and whites. COMMIT Research Group. Am J Public Health. 1993;83(2):220–6.

Sarasohn-Kahn J. The wisdom of patients: health care meets online social media [Internet]. California HealthCare Foundation Oakland, CA; 2008 [cited 2016 Jun 24]. http://www.chcf.org/~/media/MEDIA%20LIBRARY%20Files/PDF/PDF%20H/PDF%20HealthCareSocialMedia.pdf

Schvaneveldt RW. Pathfinder associative networks: studies in knowledge organization [Internet]. Ablex Publishing; 1990 [cited 2016 Jun 24]. http://psycnet.apa.org/psycinfo/1990-97976-000
Schvaneveldt, Roger, Cohen, Trevor. Abductive reasoning and similarity. In: Ifenthaler D, Seel NM, editors. Computer based diagnostics and systematic analysis of knowledge. New York: Springer; 2010.
Shalizi CR. Comment on Why and when "flawed" social network analyses still yield valid tests of no contagion. Stat Politics Policy 2012;3(1):5.
Shalizi CR, Thomas AC. Homophily and contagion are generically confounded in observational social network studies. Sociol Methods Res. 2011;40(2):211–39.
Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498–504.
Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol. 2008;4:1–32.
Smith CA, Wicks PJ. PatientsLikeMe: consumer health vocabulary as a folksonomy. In: AMIA annual symposium proceedings [Internet]. American Medical Informatics Association; 2008 [cited 2014 Mar 24]. p. 682. http://www.ncbi.nlm.nih.gov/pmc/articles/pmc2656083/
Sridharan V, Cohen T, Cobb N, Myneni S. Characterization of temporal semantic shifts of peer-to-peer communication in a health-related online community: implications for data-driven health promotion. Paper to be presented at AMIA Annual Symposium Proceedings 2016, Chicago, IL.
Strauss A, Corbin J. Basics of qualitative research: techniques and procedures for developing grounded theory. 2nd edn. Thousand Oaks, CA: Sage Publications Inc.; 1998. Xiii 312 pp.
Tang X, Yang CC. Identifying influential users in an online healthcare social network. In: Intelligence and Security Informatics (ISI), 2010 IEEE International Conference on [Internet]. IEEE; 2010 [cited 2016 Jun 24]. p. 43–8. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5484779
Tang J, Abraham C, Stamp E, Greaves C. How can weight-loss app designers' best engage and support users? A qualitative investigation. Br J Health Psychol. 2015;20(1):151–71.
Turney PD, Pantel P, et al. From frequency to meaning: vector space models of semantics. J Artif Intell Res. 2010;37(1):141–88.
Valente TWTW. Network models of the diffusion of innovations [Internet]. 1995 [cited 2016 Jun 25]. http://www.sidalc.net/cgi-bin/wxis.exe/?IsisScript=sibe01.xis&formato=2&cantidad=1&expresion=mfn=019132
Umberson D, Montez JK. Social relationships and health a flashpoint for health policy. J Health Soc Behav. 2010;51(1 suppl):S54–66.
Urberg KA, Değirmencioğlu SM, Pilgrim C. Close friend and group influence on adolescent cigarette smoking and alcohol use. Dev Psychol. 1997;33(5):834.
Valente TW. Network models and methods for studying the diffusion of innovations. Models Methods Soc Netw Anal. 2005;28:98.
Valente TW. Social networks and health: models, methods, and applications. Oxford University Press; 2010 [cited 2014].
Valente TW. Network interventions. Science. 2012;337(6090):49–53.
Valente TW, Pumpuang P. Identifying opinion leaders to promote behavior change. Health Educ Behav. 2007;34(6):881–96.
Valente TW, Gallaher P, Mouttapa M. Using social networks to understand and prevent substance use: a transdisciplinary perspective. Subst Use Misuse. 2004;39(10–12):1685–712.
Valente TW, Fujimoto K, Chou C-P, Spruijt-Metz D. Adolescent affiliations and adiposity: a social network analysis of friendships and obesity. J Adolesc Health. 2009;45(2):202–4.
Valente TW, Fujimoto K, Palmer P, Tanjasiri SP. A network assessment of community-based participatory research: linking communities and universities to reduce cancer disparities. Am J Public Health. 2010;100(7):1319–25.
Van De Belt TH, Engelen LJ, Berben SA, Schoonhoven L. Definition of Health 2.0 and Medicine 2.0: a systematic review. J Med Internet Res. 2010;12(2):1–14.

VanderWeele TJ, Ogburn EL, Tchetgen Tchetgen EJ. Why and when "flawed" social network analyses still yield valid tests of no contagion. Stat Polit Policy [Internet]. 2012 [cited 2016 Jun 25];3(1). http://www.degruyter.com/view/j/spp.2012.3.issue-1/2151-7509.1050/2151-7509.1050.xml

Vidya Vasuki, Cohen, Trevor. Reflective random indexing for semi-automated indexing of MEDLINE abstracts. Journal of Biomedical Informatics. 2010;44(2):240–256.

Velardi P, Navigli R, Cucchiarelli A, D'Antonio F. A new content-based model for social network analysis. In: Semantic Computing, 2008 IEEE International Conference on [Internet]. IEEE; 2008 [cited 2014 Mar 24]. p. 18–25. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4597169

Wang P, Robins G, Pattison P, Lazega E. Exponential random graph models for multilevel networks. Soc Netw. 2013;35(1):96–115.

Wasserman S, Pattison P. Logit models and logistic regressions for social networks: I. An introduction to Markov graphs and*p*. Psychometrika. 1996;61(3):401–25.

Widdows D, Cohen T. Graded semantic vectors: an approach to representing graded quantities in generalized quantum models. In: Atmanspacher H, Filk T, Pothos E, editors. Quantum interaction. Lecture Notes in Computer Science. 9335. Springer International Publishing; 2016. p. 231–44.

Widdows D, Cohen T. The Semantic Vectors Package. New Algorithms and Public Tools for Distributional Semantics. Semantic Computing (ICSC), 2010 IEEE. Fourth International Conference on. p 9–15.

Widdows D, Ferraro K. Semantic vectors: a scalable open source package and online technology management application. In: Sixth International Conference on Language Resources and Evaluation, LREC; 2008.

Wipfli HL, Fujimoto K, Valente TW. Global tobacco control diffusion: the case of the framework convention on tobacco control. Am J Public Health. 2010;100(7):1260–6.

Zhang M, Yang CC, Gong X. Social support and exchange patterns in an online smoking cessation intervention program. In: Healthcare Informatics (ICHI), 2013 IEEE International Conference on [Internet]. IEEE; 2013 [cited 2014 Mar 24]. p. 219–228. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6680481

Zhang S, Grave E, Sklar E, Elhadad N. Longitudinal analysis of discussion topics in an online breast cancer community using convolutional neural networks. ArXiv Prepr ArXiv160308458 [Internet]. 2016 [cited 2016 Jun 24]; http://arxiv.org/abs/1603.08458