An Appraisal of Social Network Theory and Analysis as Applied to Public Health: Challenges and Opportunities

Thomas W. Valente and Stephanie R. Pitts

Institute for Prevention Research, Department of Preventive Medicine, School of Medicine, University of Southern California, Los Angeles, California 90034; email: tvalente@usc.edu

Annu. Rev. Public Health 2017. 38:103–18

First published online as a Review in Advance on December 15, 2016

The *Annual Review of Public Health* is online at publhealth.annualreviews.org

https://doi.org/10.1146/annurev-publhealth-031816-044528

Keywords

social network analysis, social influence, social media, intervention evaluation

Abstract

The use of social network theory and analysis methods as applied to public health has expanded greatly in the past decade, yielding a significant academic literature that spans almost every conceivable health issue. This review identifies several important theoretical challenges that confront the field but also provides opportunities for new research. These challenges include (*a*) measuring network influences, (*b*) identifying appropriate influence mechanisms, (*c*) the impact of social media and computerized communications, (*d*) the role of networks in evaluating public health interventions, and (*e*) ethics. Next steps for the field are outlined and the need for funding is emphasized. Recently developed network analysis techniques, technological innovations in communication, and changes in theoretical perspectives to include a focus on social and environmental behavioral influences have created opportunities for new theory and ever broader application of social networks to public health topics.

INTRODUCTION

The aim of this review is to assess the current state of the science of applying social network theory and analysis (SNT/A) in the field of public health. It is not intended as an introduction to SNT/A, and interested readers may consult other introductions (7, 43, 55, 79, 81). Rather we attempt to address the challenges and outstanding research questions that confront the field today. We also do not attempt to review all the current research being conducted to apply SNT/A to public health topics as the number of research fronts currently being explored is indeed rather large. Instead, this article is designed to train a critical eye on several scientific issues that researchers have yet to resolve and to provide opportunities for new research.

SNT/A has been applied to almost every area of health research, including physician behavior (15, 29, 41), adolescent risk taking (30), obesity and physical activity (16), bullying (62), community-based participatory research (85), policy diffusion (84), interorganizational relations (25, 49, 68), and community coalitions (82) among other things. The methods and settings have also been quite varied, including social support scales (36), egocentric studies (64), snowball-type studies (32), organizational ones (93), and studies in both developed- and developing-country contexts (27, 46). Thus, we have many examples from which to illustrate challenges and opportunities facing network researchers.

Network research in public health classically focused on the transmission of infectious diseases. For millennia, newly established trade, migration, and communication networks connected previously disconnected communities, resulting in cataclysmic transmission of deadly viruses (17, 58). By the twentieth century, public health officials employed tools such as contact tracing in attempts to treat sexually transmitted infections and other diseases. Formal application of network concepts, however, was rare. The HIV/AIDS epidemic in the late 1980s challenged public health researchers to understand how contact networks, especially syringe sharing and/or sexual relations, influenced HIV spread (61). Network research subsequently transitioned from a focus on infectious diseases to chronic diseases. Although quite a few behavioral studies had been conducted prior to the 1990s, it was at this time that growth accelerated in the application of social network analysis to behaviors associated with chronic conditions.

By the time of the influential study by Christakis & Fowler (12), the transition was complete, and today SNA is applied much more frequently to behaviors associated with chronic conditions than to infectious ones. These application areas include tobacco, alcohol, and other drug use; obesity and physical activity; reproductive health; health services; and many other areas. The focus on behaviors rather than diseases stems in part from available data; few if any data sets contain contact information and disease spread, though see Vanhems et al. (91). It also stems from the seriousness of chronic conditions because these now constitute the major public health risk factors in most populations.

The sine qua nons of network research are the collection and analysis of network data, that is, information on the connections and relationships among and between entities. Usually, though not exclusively, these entities are people such as the general population or subgroups such as adolescents in schools, members of an online community, and organizational employees. In what follows, our remarks are meant to signify studies among people and usually in a somewhat bounded community, though the size of such boundaries continues to increase owing to advances in computing technology and the use of online or social networking media. We recognize, however, that many health network studies are conducted among organizations, states, nations, and/or other units. Our review of network research has identified four broad areas that represent challenges to the continued advance of network applications in public health: (*a*) estimating network influences or so-called contagion effects, (*b*) articulating appropriate theoretical mechanisms by which networks have effects, (*c*) disentangling mediated versus nonmediated network processes, and (*d*) the application of network theory and methods for intervention design and evaluation. We address each of these as challenges and opportunities.

NETWORK INFLUENCE

A rather large community of scientists has been active for decades, developing theories, algorithms, and metrics to describe social networks (24). These scholars have produced measures of centrality; brokerage and bridging; group and community detection; blockmodeling; positional analyses; and network-level metrics such as density, clustering, average path length, and network diameter among others. Interested scholars can take workshops, courses, and tutorials to learn the math, techniques, and software necessary for network analysis (e.g., **http://www.insna.org**). Distinct from this line of work, however, many researchers have focused on trying to determine whether and how social networks influence health behavior.

The earliest of these studies were based on diffusion of innovations theory, which emerged in the early 1940s as a distinct research paradigm (88). The classic medical innovation study by Coleman and others (15) is often considered the first to attempt to measure social networks and their association with innovation adoption. Several network diffusion studies followed, and many studies were conducted to reanalyze data from the three classic remaining diffusion network data sets (18, 76, 79). Diffusion, however, takes a long time; diffusion data have traditionally been difficult to obtain, so many researchers focused on network influences on short-term behavior change (3-, 6-, or 12-month follow-up).

As a practical matter, many early studies were cross-sectional and associational. For example, Valente and others (90) showed that Cameroonian women were more likely to be aware of and use the same contraceptive methods as their friends. Moreover, these women were more likely to use contraceptive methods when they perceived their friends encouraged contraceptive use. So even though the data were cross-sectional, the logic and theory supporting the analysis lent evidence to a network effect. Other studies of network influences on contraceptive use have been conducted and also show an association between use and peer use (20, 27, 70).

The application area with perhaps the most evidence of network influences is adolescent tobacco use. Despite progress in reducing adolescent tobacco use, it is still the leading cause of premature mortality globally, and 80% of adult smokers report initiating tobacco use when they were adolescents (63). Researchers have hypothesized for many decades about the existence of a network effect on initiation and continuation of tobacco use (19). Using the Add Health data, Alexander and colleagues (1) showed that smokers were more likely to have smoking friends, popular students were more likely to be smokers, and they were more likely to be smokers in schools with a high smoking prevalence.

Ennett & Bauman (19) published a landmark study postulating both influence and selection mechanisms as explanations for homophily on tobacco use. Influence occurs when individuals make behavior changes to be consistent with their peer network. Conversely, selection occurs when individuals make network changes to be consistent with their behavior. Selection provided an alternative (to influence) explanation for behavioral homophily. The debate over selection versus influence has intensified in recent years, particularly with the advent of stochastic actor-oriented models (SAOM), which can compare selection and influence mechanisms simultaneously (59), though selection can be tested without using SAOM (86). The challenge is that SAOM have more power to detect selection than to influence, which may bias the interpretation of results (we return to this point below).

In sum, many studies have shown that networks are associated with adolescent tobacco use, contraceptive choices in developing countries, and other behaviors. However, making causal claims about network effects is fraught with statistical challenges (72). First, network data are often nonindependent, which violates a basic assumption of inferential statistical analysis. Second, there is endogenous tie formation: A person becomes friends with the friends of his/her friends. If homophily on smoking exists, then these new ties will seem like influence effects but are really a function of natural network evolution. Third, there is latent homophily, that is, people are connected for reasons unobserved in the study yet may be associated with the behavior. For example, two people may be friends because they belong to a club outside the school, and if smoking becomes prevalent in that club it will be mistaken for influence from their school networks.

Consequently, the many studies that show an association between behavior and peer (network) behavior are often observational and fail to control for these threats to interpretation. There is reason for optimism, however, and these challenges provide new opportunities for research that specifically address these shortcomings.

Researchers should continue to push for collecting longitudinal data on networks and behavior and for additional measures of the social environment. By collecting data over many time points, greater specificity of the theoretical mechanisms that drive behavior change can be made and used to address the concerns outlined above. Moody and others (60) provide an excellent example of longitudinal network associations with behavior that control for many alternative explanations. Latent homophily may explain associations between behavior and network at baseline or in the short term; however, over long periods of time, the effects of homophily diminish as influence mechanisms kick in. Furthermore, supplementary data on the broader social environment characteristics may be important for understanding how the structure and characteristics of social networks affect behavior change and for controlling for potential causes of homophily in the network.

Longitudinal data may also be analyzed using the SAOM approach, which can control for demographic homophily and endogenous tie formation in the analysis. Although not every potential latent homophily variable can be measured or envisioned, the main ones, which provide an alternative explanation for influence, can be. For example, among adolescents, gender may influence both tie formation and risk behavior, and SAOM as implemented in SIENA enables the researcher to control for homophily on gender while testing for homophily on behavior. Multilevel models (96) can be used to incorporate multiple individual attributes and networks in analysis of network influences on behavior change.

Diffusion studies typically consist of many time intervals over which behavior change occurs, thus providing a longitudinal framework within which to test for network influences (84). The added advantage of diffusion research is that it provides a rich theoretical history spanning more than a century, which investigators can use in the specification of network effects (18, 88). Diffusion data lend themselves to survival analysis and event history analysis as well as many unique metrics that capture peer influences and network effects on adoption (T.W. Valente, submitted).

Experiments also provide a way to test for network influences (11). Early network research was characterized by experiments (3), and researchers can continue this tradition by conducting experiments specifically with network interventions (80) or by manipulating network conditions in online environments as others have done (5, 11). Field experiments (46, 54) have generally documented strong network effects, and they build support for network theory.

There is clearly much work to be done to solve the influence/causality issue, which will not be resolved with one study or one finding. The continued accumulation of evidence regarding how networks influence, and are influenced by, behavior will enable us to articulate the context and conditions under which these effects occur. From a policy perspective, it is important to know how social networks influence behavior so that proper attribution can be made. Like most scientific endeavors, theory will help guide the way, and it is to theory we now turn to discuss the specification of the causal mechanisms by which networks have their effects.

MECHANISMS

Out of convenience, the preceding discussion has assumed that network influence occurs when an individual is exposed to other individuals who engage in a particular behavior and who persuade the individual to adopt the new behavior. That exposure can result in behavior change via several different mechanisms such as persuasion, communication of norms, modeling, information, support, social pressure, etc.

Yet little network research has focused on or attempted to contrast different mechanisms of how networks influence behavior. The seminal study by Burt (9) proposed that network influence can occur via structural equivalence, that is, individuals adopt innovations when they see others who are equivalent to them, in network terms, adopt the innovation. The mechanism of network influence was thus conceived not as communication and the sharing of information but rather as competition or the pressure individuals feel to conform to behavior when others in equivalent positions do so. Yet cohesion and structural equivalence are but two of the many mechanisms by which networks can have effects (T.W. Valente, submitted).

Network thresholds were another mechanism discovered to explain network effects (77). Network thresholds are the number or proportion of people in one's network who must adopt the behavior before an individual is willing to do so. People with low thresholds adopt behaviors before any or few of their peers do so. Thresholds provided the means to test the famous two-step flow hypothesis (44), which stated that the media influenced opinion leaders who, in turn, influenced others in their networks. The two-step flow model with thresholds specified that the media (or other external communications such as cosmopoliteness) influenced low-threshold adopters because they did not have adopters in their network to turn to for information about the behavior (89); these low-threshold people then persuaded others.

Other mechanisms include network-weighted exposures. For example, we might hypothesize that opinion leaders have a greater influence on subsequent diffusion than do other people in the community. Consequently, exposures may be weighted by the in-degree scores in the network so that when opinion leaders adopt, they influence their peers more than when nonleaders adopt. Attribute-weighted exposures may also be tested such that adoptions by high-SES (socioeconomic status) network members could have greater influence than would low-SES ones.

Developing and applying theories that articulate how different influence mechanisms act under different conditions provide a means to specify network influences without creating steep data requirements. Intelligent application of network theory may obviate the need for massive amounts of data or laboratory experiments. Studies comparing influence mechanisms provide a gaping opportunity for new network research.

In addition, the advent of new communication technologies has expanded the realms in which interpersonal influences can occur. Persuasion may result even when the persuader does not intend to persuade. For example, people see their friends engaging in behaviors through their online postings, which may influence a person's behavior regardless of whether the poster intended to do so. Social learning occurs, and there is a data trace to capture it, but the mechanism of change is quite different than in-person purposive persuasion. Even in person, interpersonal influence may occur through the effects that perceptions of social interactions have on people's views of themselves and how that may consequently affect their behavior (95).

The variety of network influence mechanisms shifts the locus of behavior change from cognitive or individual explanations to social ones. The history of understanding behavior has evolved from an emphasis on biological and psychological factors that govern behavior to an emphasis on contextual and social ones. The contrast between these two orientations, individual versus social, is apparent in two of the most widely used stages of change models: the transtheoretical model, which emphasizes cognition (premeditated, contemplative, preparation, action, maintenance), and diffusion of innovations (awareness, persuasion, decision, trial, adoption), which emphasizes sources of information and influence. The transition from cognitive factors to social ones broadens the number of concepts and mechanisms that can be incorporated into behavior change models and expands the sources of data used to answer these questions. Furthermore, it opens up opportunities for theory development focused on the inclusion of multiple levels of factors that influence behavior and the articulation of how factors on different levels influence one another (71). The best research, of course, will combine concepts and tools from all disciplines and incorporate measures from cellular to broad systems and policies.

INTERNET MEDIATED

Certainly one of the most exciting developments in health applications of social network analysis is the capability to mine social media and other electronic sources of data (51). Such data address two biting criticisms leveled at early network research: (*a*) that it was based on small idiosyncratic samples of dozens of people (rural villages, schools, organizations, etc.); and (*b*) it was based on either survey or observation data and therefore prone to both missing data and error. Social media and other platforms automatically record network connections, thus providing accurate network data, and can do so on a large scale, which makes network analysis scalable.

Several implications are inherent in the step from small-scale studies to larger mediated ones. First, can influence occur via mediated communications to the extent that it does in face-to-face interaction? The evidence suggests that mediated communications can influence individual behaviors but they do so at a rate much lower than face-to-face communications do (37). Still, people can have hundreds, even thousands or more, of online contacts, which increases the potential for much influence. Moreover, some forms of mediated communications may be particularly influential. For example, specialty communities that emerge in forums such as PatientsLikeMe may be very influential because they offer a place where people can share information about extremely important and relevant topics among members of a specific community. For example, in our recent study (66), researchers showed that individuals' concerns about their weight and appearance are significantly influenced by both their online and offline social networks and that this influence varies by body-mass index.

Second, comparing studies based on small, in-depth analyses of communities with those conducted among large populations provides the opportunity to expand and deepen network theory. Studies conducted in rural villages or schools often collect lots of rich data through survey, observation, or both. These data can be used to explore theoretically important hypotheses regarding human behavior and health. Conversely, substantial amounts of social media data may have only limited data on the individuals themselves. Twitter data, for example, provide extensive information on links between Twitter users. However, data on the characteristics of the users and their motivations for tweeting are lacking (57). Researchers are currently working to rectify the paucity of data on individuals by developing methodologies for determining the demographics of social media users, such as coding demographics from profile photos (57). Still, the smaller studies should be able to provide theory that informs the larger ones.

Table 1 contrasts big data studies, which often acquire data from computer-mediated communications, with depth data, which often acquire data from surveys and/or observation. No studies, to our knowledge, have followed through on this continuum, taking ideas developed in the lab, to

Table 1 Contrasting big data, depth data, and simulated data

	Big data	Depth data	Simulation
Examples	Twitter, Facebook, Instagram, recommender systems	Students in schools, organizations, rural villages	Susceptible, exposed, infected, recovered models
Advantages	Large sample size Results apply to many people	In-depth data on attributes and links Easier to manage	Do not need to collect data Can model hypothetical scenarios
Disadvantages	Little attribute information Restricted by technology	Less generalizable Missing data	May not accurately simulate dynamics occurring in the real world
Example Studies	Mobile network data used to track probable disease spread (92)	Economic and social networks of villages created with consultation of village elders. Qualitative data collected on villager interactions with Ebola victims (69)	Cases and deaths predicted from Ebola using a network-based simulation (74)
What did the studies provide?	Describes population movement that is relevant for policies on movement restriction	Explains why people travel often between certain villages and why people came in contact with one another, providing context for policy	Provides predictions that are useful for planning health care and prevention services
Analytic tools	R, C^{2+}, Python	R, UCINET, Pajek	Matlab, Netlogo, R
How acquired	Online, sensors, phones, preexisting databases	Paper-and-pencil surveys, interview, data collection software	User generated

small-scale studies, to large scaled-up ones, but such studies would be a promising development. The lack of studies that cross scales suggests the need to conduct a study of network influence, which applies the same concepts and process in silico, as computer simulation; in vitro, as laboratory or clinical experimentation; and in vivo, in the real world, and ideally progress from depth to big data. We have found no studies that have explicitly compared network data gathered online with data gathered offline, although this approach seems like a good idea.

Third, not all mediated communications are the same. Social media involves the sharing of personal information with preselected friends, family, or colleagues. Referral systems, which recommend products to users on the basis of past product preferences and comparison to others' preferences, are quite different and entail communications about the behaviors of others who are, more than likely, strangers and unknown to the focal individual. Community sites gather people with similar interests or concerns and provide forums for information, exchange, and support (33). The influences that occur on these three, and other, types of mediated communication can be very different. The affordances inherent in various media and platforms may have implications for different network influences (13). Studies are needed to compare how individuals react to and use mediated communications that influence their health across various forms of media, acknowledging that data from many electronic sources may not be representative. These observations suggest that policy makers need to monitor social media activity across a wide spectrum to understand how populations and subgroups feel about public health issues.

EVALUATIONS

An emerging application of SNT/A is in the creation, design, and implementation of network interventions to promote health (80, 87). Networks are also significant factors that affect the evaluation of all public health interventions in at least three ways: (*a*) estimating spillover or contagion effects, (*b*) understanding true prevalence, and (*c*) measuring moderation and mediation of program effects. Spillover effects occur when program participants communicate their experiences to others, either directly or indirectly influencing behavior change in comparison group populations or among the general population (65, 67). Programs become more effective when participants communicate with their peers and spread the word about favorable behavior changes from an intervention. These spillover effects can be captured using SNA techniques and will thus generate more accurate, and more positive, estimates of program effects.

Not all community members are equal from a network perspective. **Figure 1** illustrates a hypothetical situation with a network of 30 people and initial behavior prevalence of 13.3% (4 of 30) and 16.7% (5 of 30). These initial users could be central members of the network or peripheral ones. If they are the most central nodes, as in **Figure 1*a***, subsequent diffusion of the behavior is likely to be quick, as everyone is connected to at least one existing user; overall these central nodes provide an average rate of exposure to others of 34.5%. In contrast, if the initial users are on the periphery, as in **Figure 1*b***, then diffusion is likely to be slow if at all. These peripheral nodes do not expose anyone else to the innovation, so there is no potential for diffusion. So instead of two communities with equal prevalence, we have one community (**Figure 1*a***), in which many ties in the network are to existing users, thus providing influence or reinforcement for the new behavior. In the other community (**Figure 1*b***), there are few outgoing ties to users (0%) and so no opportunity for influence or reinforcement. This phenomenon has been referred to as the majority illusion and explains how a behavior that is uncommon may be viewed as normative owing to social network structure (52).

Some evidence indicates that networks moderate and/or mediate intervention effects. Intervention effects moderation occurs when programs are effective for some subgroups and not others. The network threshold model is an example of effects moderation (77): One study showed that a mass media intervention was effective for audience members with low thresholds and not those with high ones (e.g., 89). People whose networks have contraceptive users can turn to these individuals for information and advice, whereas those in networks without contraceptive users must turn to other sources of information and influence. Shin and others (73) showed that a physical activity intervention reduced negative peer influences on unhealthy behaviors.

Mediation occurs when interventions are posited to change variables that, in turn, create changes in outcomes. For example, interventions designed to change perceived norms for behavior are effective when they change norms and these norms change behaviors (22). Many group-based interventions are designed to create new links among participants so that they can reinforce new healthful behaviors (28). Gesell and others (28) showed that a group-based healthy lifestyle intervention was successful at creating new advice and discussion networks in a relatively short time frame (three weeks). These network changes were also associated with changes in perceived group cohesion. It is anticipated that this increased cohesion will result in greater behavior change. Thus, intervention effects may be moderated and/or mediated by network properties, their level of effectiveness being based on (*a*) network position (e.g., central versus peripheral), (*b*) network composition (e.g., the number or proportion of users), and/or (*c*) overall network structure. Consequently, evaluation plans should incorporate network measures to validly estimate program outcomes.

ETHICS

Despite the opportunities made available by the Internet for social network research, researchers, ethicists, and the public have raised concerns about what constitutes ethical research using the Internet. Lewis and others (53) constructed a data set of Facebook profiles of students enrolled at a northeastern university and made it publicly available for research. Soon after, Zimmer (97) raised

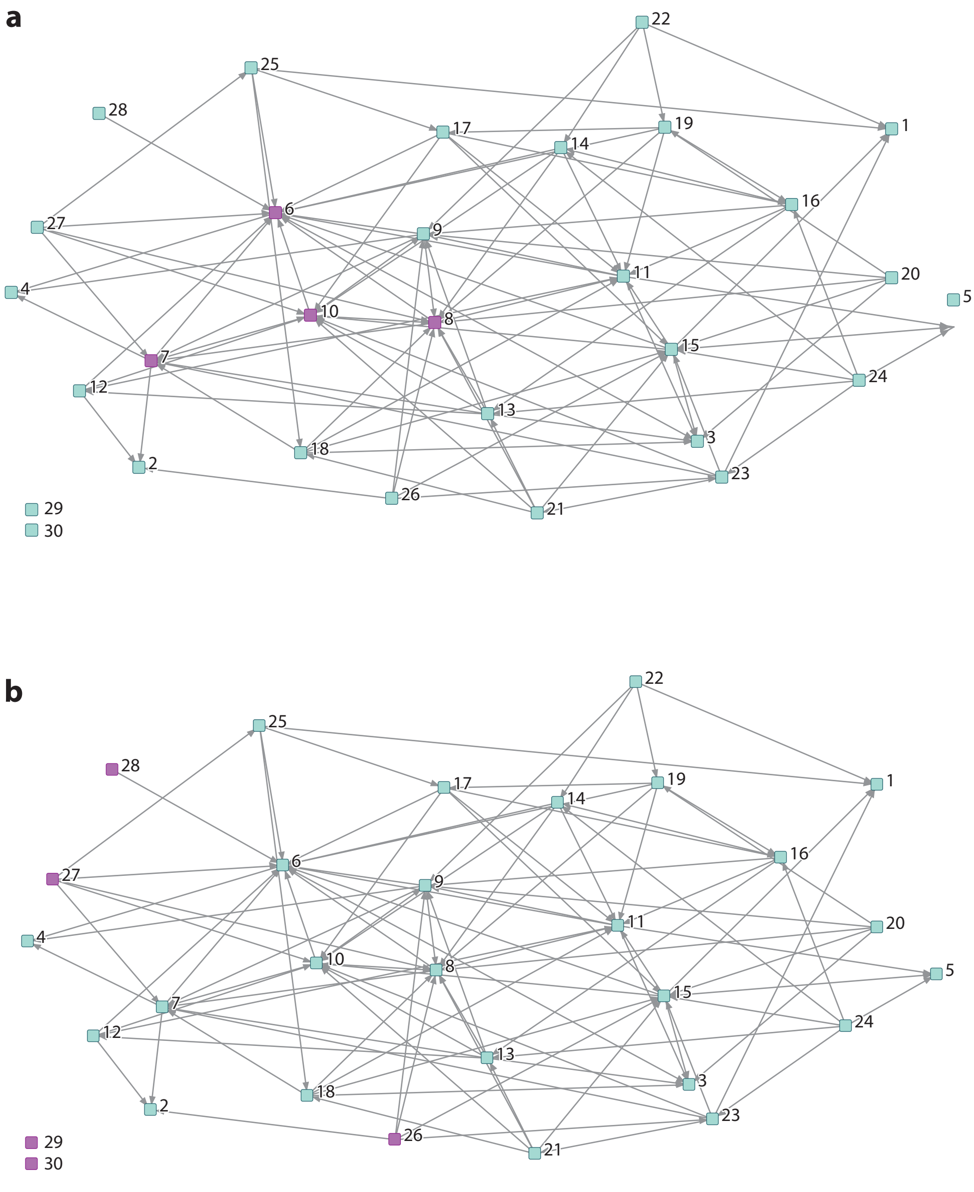

Figure 1

A randomly generated 30-node network with 4 (13.3%) users who are the most central and the same network with five (16.7%) users who are peripheral. Although both have similar initial prevalence, future diffusion in panel *a* is likely to be much faster than in panel *b*. In panel *a*, the average exposure to the innovation is 34.5%, whereas in panel *b* it is zero. This network was generated using a scale-free random graph model based on Bollabas et al. (2001). The figures were drawn using Netdraw (6).

privacy concerns about the anonymity of the data and the lack of informed consent obtained from the students in the data set. Zimmer also claimed to know which university the data came from.

In a similarly publicized study, Kramer and others (48) used Facebook as a social psychology laboratory and manipulated users' Facebook News Feeds to induce negative or positive affect to study the social contagion of emotions. Facebook users did not provide overt consent to participate in the study, and many felt that their rights were violated (35). Furthermore, participants of in-person experimental studies are often debriefed when their study participation has concluded to reduce the likelihood that the study's manipulations and procedures caused harm, yet no such debriefing was performed in this study. Mood manipulations in the study by Kramer et al. (48) may have caused harm to those people who have a mood disorder and were unaware that they had been manipulated to feel worse (35). These studies raise concerns about the anonymity of social media data, unintended use of social media data and platforms, standards for informed consent, and the unintended effects of online experimental manipulations. No standardized regulations have been developed yet for Internet research, although the Association of Internet Researchers has published a set of guidelines for Internet research (31, 56). Furthermore, researchers have worked to develop methodologies for data anonymization for publishing record-based data, which may be adopted and adapted to network research (26). In addition to these sources, researchers conducting online studies may consult publications by Innovate UK, NatCen Social Research, the British Psychological Society, the Economic and Social Research Council (ESRC), and the British Library to inform ethical considerations (4, 21, 75, 94).

Most social network data are not anonymous but should be confidential. Online and offline network data collection often includes identifying alters by their true names. Furthermore, social network surveys may contain questions about the participants' alters, which are sensitive and may elicit information that the alter would otherwise not allow researchers to have (47). Some discussion has surrounded whether alters should be considered research participants and whether informed consent should be obtained from alters (47). Social network research within public health may result in the collection of sensitive data on alters, including medical and genetic conditions, and if the information is compromised, these individuals may experience negative financial, psychological, or legal effects (47).

Similar concerns arise with data accessed on social media sites (39). The function of network analysis is to make things apparent that were not in the original presentation of the data (42). Users of social media lose control over their information when it is presented in a new format or analyzed in a way that information about them becomes unintentionally available (56, 97). Furthermore, issues of anonymity are raised when researchers turn study findings and data over to the populations involved in the study (42). Members of a network may be able to identify themselves and others in a network diagram, even if the data are anonymized (42). Although the study population has the right to benefit from participating in the study, turning over the results of social network analysis may jeopardize the rights of some or all study participants, even if the results are anonymized.

Participation in social network studies can bring about unique burdens and potential harms. Network questionnaires may require participants to respond to numerous questions per alter named, resulting in an increased burden on the participant (83). Furthermore, participation in a social network study may highlight a participant's social reality and cause stigmatization. For example, if a participant cannot think of anyone to list as a friend on a survey, they may experience sadness. New procedures for obtaining consent for research on social media sites may reduce harms associated with Internet research and allow participants more control over what they share; however, responding to researchers' requests for information may also increase the burden on study participants (40).

Potential harms due to privacy loss are not always apparent to participants. Potentially identifiable network data collected from social media can compromise a person's dignity, even if his or her data are not used in a way that could otherwise cause harm (97). Kadushin (42) states the participants of social network research rarely benefit from participation. The unique risks and burdens associated with participation in social network studies may not be readily apparent to researchers, participants, and institutional review boards. These concerns should not stop social network research but instead provide opportunities for studies that can inform the extent of, and conditions for, burden, as well as benefit, imposed by network research.

NEXT STEPS

Experience with public health interventions to date has made it clear to researchers that changing the knowledge, attitudes, and practices of selected populations is possible with well-designed and well-implemented interventions. These effects are typically modest and short-lived, commensurate with the resources applied to the problem and with acknowledgment of the considerable barriers blocking change. The challenges facing continued progress in improving public health are at least threefold: (*a*) understanding mechanisms of change, (*b*) sustaining change after the resources are withdrawn or the program ends, and (*c*) scaling interventions to have wider impacts. Understanding social networks and their deployment in behavior change programs offers the promise to address all three of these challenges.

As outlined in the first section of this article, we believe that social networks provide the means to understand mechanisms for behavior change. Most behavior change models acknowledge that people progress through a set of stages of change whether based on diffusion of innovations, the transtheoretical model, or the hierarchy of effects (78). Social networks play a pivotal role in each of the stages of change and become more important as people progress from awareness or precontemplation to the ultimate steps in the process. People learn, contemplate, acquire information, try, and ultimately adopt new behaviors in the context of their interpersonal relationships. The more complex or challenging the behavior change topic, the more people rely on their social networks at each stage of change.

Sustaining change has long been a frustration of public health interventionists (65). Quite often once a program ends, the funding for the services and the change agent training disappear. This lack of continuation signals to the community that the problem is no longer important, and therefore the energy previously devoted to the problem is directed elsewhere. For example, funding cuts to obesity-prevention programs resulted in reductions in prevention programming due to loss of staff, infrastructure, and key partners (23). By deploying social network techniques in the behavior change program, sustainment can be achieved with no additional cost. Because the change program is designed and implemented by community members who remain embedded in the community once the program is over, the change process continues and is reinforced. For example, Kelly and others (45) identified peer opinion leaders within ethnic Roma communities in Eastern Europe and trained them in techniques to promote behavior change. The three-month intervention period ended with significant reductions in HIV risk behaviors, and nine months later these changes persisted because the opinion leaders were still embedded within their communities promoting behavior change, being a role model, and providing a constant reminder of the lessons learned during the intervention. Behavior change maintenance has also been found in the ASSIST trial, a school-based peer-led study to prevent tobacco use (10).

Furthermore, it is difficult for people to sustain their own personal behavior change, and many people relapse to their original behaviors after completing a behavior change intervention (50). In a review of theories of behavior maintenance, Kwasnicka and others (50) identified environmental and social influences as important theoretical themes and stated that behavioral maintenance is influenced by the social environment and social support. A recent study by Hunter and others (38) found that in a workplace physical activity intervention, social networks existed that influenced behavior change and maintenance, despite networks not being included in the intervention design. Social networks may assist or deter a person from sustaining positive health changes and should be considered when designing interventions.

Scaling up interventions is also a challenge. Social network interventions have received criticism for their lack of generalizability, but we would counter that this limitation is due only to a lack of research guiding the scaling-up of social network interventions. Furthermore, how can we take an intervention that works in one community and expand it to have regional or national reach, thus extending the cost-effectiveness of the intervention? Our best hope to achieve this goal is through social media and the new computational tools becoming available to social scientists (8). The Web became a prominent locus of interaction when interactivity became incorporated into it. Research has shown that people can be influenced by messages and postings to which they are exposed online, and some experiments have been conducted to show that information and behaviors can be spread online (5, 13, 14, 34). Large-scale public health online campaigns have not yet produced favorable results, however, and at least one instance triggered a countercampaign (2). Figuring out the right way to deploy social media in the service of public health campaigns clearly presents a significant challenge and opportunity going forward.

Much work remains to be done to advance the science of network studies in public health. First, experiments, both randomized and quasi-experimental, are needed to identify network influences in ways that rule out endogeneity and other contaminating effects. Longitudinal studies that compare different networks and different network mechanisms are needed to specify more fully the conditions under which influence occurs and why.

None of this work can take place without funding from the National Institutes of Health and leading foundations. The most significant scientific advances depend on funding to collect data and develop the computational tools necessary to make the comparisons called for in this review. New work on network interventions (46, 80) promises to address the challenges faced by public health interventionists, advocates, and researchers.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This study was supported by grant R01 CA157577 from the National Cancer Institute. Our thanks go to Kar-Hai Chu, Rebecca Davis, Ruth F. Hunter, Jieun Shin, and George G. Vega Yon for comments on earlier drafts and members of the Center for Applied Network Analysis at the University of Southern California.

LITERATURE CITED

1. Alexander C, Piazza M, Mekos D, Valente T. 2001. Peers, schools, and adolescent cigarette smoking. *J. Adolesc. Health* 29(1):22–30
2. Allem J-P, Escobedo P, Chu K-H, Soto DW, Cruz TB, Unger JB. 2016. Campaigns and counter campaigns: reactions on Twitter to e-cigarette education. *Tob. Control* **https://doi.org/10.1136/tobaccocontrol-2015-052757**

3. Bavelas A. 1950. Communication patterns in task-oriented groups. *J. Acoust. Soc. Am.* 22(6):725–30
4. Beninger K, Fry A, Jago N, Lepps H, Nass L, Silvester H. 2014. *Research Using Social Media; Users' Views*. London: NatCen Soc. Res.
5. Bond RM, Fariss CJ, Jones JJ, Kramer ADI, Marlow C, et al. 2012. A 61-million-person experiment in social influence and political mobilization. *Nature* 489:295–98
6. Borgatti SP. 2002. *NetDraw Software for Network Visualization*. Lexington, KY: Anal. Technol. **https://sites.google.com/site/netdrawsoftware/home**
7. Borgatti SP, Everett MG, Johnson JC. 2013. *Analyzing Social Networks*. Los Angeles: Sage
8. Brown CH, Mohr DC, Gallo CG, Mader C, Palinkas L, et al. 2013. A computational future for preventing HIV in minority communities: how advanced technology can improve implementation of effective programs. *J. Acquir. Immune Defic. Syndr.* 63:S72–84
9. Burt RS. 1987. Social contagion and innovation: cohesion versus structural equivalence. *Am. J. Sociol.* 92(6):1287–335
10. Campbell R, Starkey F, Holliday J, Audrey S, Bloor M, et al. 2008. An informal school-based peer-led intervention for smoking prevention in adolescence (assist): a cluster randomised trial. *Lancet* 371:1595–602
11. Centola D. 2010. The spread of behavior in an online social network experiment. *Science* 329:1194–97
12. Christakis NA, Fowler JH. 2007. The spread of obesity in a large social network over 32 years. *N. Engl. J. Med.* 357(4):370–79
13. Chu K-H, Sidhu AK, Valente TW. 2015. Electronic cigarette marketing online: a multi-site, multi-product comparison. *JMIR Public Health Surveill.* 1(2):e11
14. Cobb NK, Jacobs MA, Saul J, Wileyto EP, Graham AL. 2014. Diffusion of an evidence-based smoking cessation intervention through Facebook: a randomised controlled trial study protocol. *BMJ Open* 4(1):e004089
15. Coleman JS, Katz E, Menzel H. 1966. *Medical Innovation: A Diffusion Study*. New York: Bobbs-Merrill
16. de la Haye K, Robins G, Mohr P, Wilson C. 2010. Obesity-related behaviors in adolescent friendship networks. *Soc. Netw.* 32(3):161–67
17. Diamond JM. 1999. *Guns, Germs, and Steel: The Fates of Human Societies*. New York: Norton
18. Dyal SR. 2016. Network influences on behavior: a summary of Tom Valente's keynote address at Sunbelt xxxv: the annual meeting of the International Network for Social Network Analysis. *Connections* 35(2):52–57
19. Ennett ST, Bauman KE. 1993. Peer group structure and adolescent cigarette smoking: a social network analysis. *J. Health Soc. Behav.* 34(3):226–36
20. Entwisle B, Rindfuss RR, Guilkey DK, Chamratrithirong A, Curran SR, Sawangdee Y. 1996. Community and contraceptive choice in rural Thailand: a case study of Nang Rong. *Demography*. 33(1):1–11
21. Evans H, Ginnis S, Bartlett J. 2015. *#SocialEthics: A Guide to Embedding Ethics in Social Media Research*. London: Wisdom of the Crowd
22. Foxcroft DR, Moreira MT, Almeida Santimano NM, Smith LA. 2015. Social norms information for alcohol misuse in university and college students. *Cochrane Database Syst. Rev.* 1:CD006748
23. Freedman AM, Kuester SA, Jernigan J. 2013. Evaluating public health resources: What happens when funding disappears? *Prev. Chronic. Dis.* 10:E190
24. Freeman LC. 2004. *The Development of Social Network Analysis: A Study in the Sociology of Science*. Vancouver, BC/North Charleston, SC: Empirical Press/BookSurge
25. Fujimoto K, Valente TW, Pentz MA. 2009. Network structural influences on the adoption of evidence-based prevention in communities. *J. Community Psychol.* 37(7):830–45
26. Fung BCM, Wang K, Chen R, Yu PS. 2010. Privacy-preserving data publishing: a survey of recent developments. *ACM Comput. Surv.* 42(4):1–53
27. Gayen K, Raeside R. 2010. Social networks and contraception practice of women in rural Bangladesh. *Soc. Sci. Med.* 71(9):1584–92
28. Gesell SB, Barkin SL, Sommer EC, Thompson JR, Valente TW. 2016. Increases in network ties are associated with increased cohesion among intervention participants. *Health Educ. Behav.* 43(2):208–16

29. Gross CP, Cruz-Correa M, Canto MI, McNeil-Solis C, Valente TW, Powe NR. 2002. The adoption of ablation therapy for Barrett's esophagus: a cohort study of gastroenterologists. *Am. J. Gastroenterol.* 97(2):279–86
30. Hall JA, Valente TW. 2007. Adolescent smoking networks: the effects of influence and selection on future smoking. *Addict. Behav.* 32(12):3054–59
31. Harriman S, Patel J. 2014. The ethics and editorial challenges of Internet-based research. *BMC Med.* 12(1):124
32. Heckathorn DD. 1997. Respondent-driven sampling: a new approach to the study of hidden populations. *Soc. Probl.* 44(2):174–99
33. Hether HJ, Murphy ST, Valente TW. 2016. A social network analysis of supportive interactions on prenatal sites. *Digit. Health* 2:1–12
34. Hinz O, Skiera B, Barrot C, Becker JU. 2011. Seeding strategies for viral marketing: an empirical comparison. *J. Mark.* 75(6):55–71
35. Holmes L. 2014. Lab rats, one and all: that unsettling Facebook experiment. *Monkey See Pop-Culture News and Analysis from NPR*, June 30. **http://www.npr.org/sections/monkeysee/2014/06/30/326923945/lab-rats-one-and-all-that-unsettling-facebook-experiment**
36. House JS, Umberson D, Landis KR. 1988. Structures and processes of social support. *Annu. Rev. Sociol.* 14:293–318
37. Huang GC, Unger JB, Soto D, Fujimoto K, Pentz MA, et al. 2014. Peer influences: the impact of online and offline friendship networks on adolescent smoking and alcohol use. *J. Adolesc. Health* 54(5):508–14
38. Hunter RF, McAneney H, Davis M, Tully MA, Valente TW, Kee F. 2015. "Hidden" social networks in behavior change interventions. *Am. J. Public Health* 105(3):513–16
39. Hutton L, Henderson T. 2013. An architecture for ethical and privacy-sensitive social network experiments. *ACM SIGMETRICS Perform. Eval. Rev.* 40(4):90–95
40. Hutton L, Henderson T. 2015. "I didn't sign up for this!": informed consent in social network research. In *Proc. Int. AAAI Conf. Web Soc. Media, 9th, Oxford*, pp. 178–87. Palo Alto, CA: Assoc. Adv. Artif. Intell.
41. Iyengar R, Van den Bulte C, Valente TW. 2011. Opinion leadership and social contagion in new product diffusion. *Mark. Sci.* 30(2):195–212
42. Kadushin C. 2005. Who benefits from network analysis: ethics of social network research. *Soc. Netw.* 27(2):139–53
43. Kadushin C. 2012. *Understanding Social Networks: Theories, Concepts, and Findings*. New York: Oxford Univ. Press
44. Katz E. 1957. The two-step flow of communication: an up-to-date report on an hypothesis. *Public Opin. Q.* 21:61–78
45. Kelly JA, Amirkhanian YA, Kabakchieva E, Vassileva S, Vassilev B, et al. 2006. Prevention of HIV and sexually transmitted diseases in high risk social networks of young Roma (Gypsy) men in Bulgaria: randomised controlled trial. *BMJ* 333:1098–98
46. Kim DA, Hwong AR, Stafford D, Hughes DA, O'Malley AJ, et al. 2015. Social network targeting to maximise population behaviour change: a cluster randomised controlled trial. *Lancet* 386:145–53
47. Klovdahl AS. 2005. Social network research and human subjects protection: towards more effective infectious disease control. *Soc. Netw.* 27(2):119–37
48. Kramer ADI, Guillory JE, Hancock JT. 2014. Experimental evidence of massive-scale emotional contagion through social networks. *PNAS* 111(24):8788–90
49. Kwait J, Marconi K, Helitzer D, Rodieck MW, Randall T, Celentano D. 1999. Ryan White Care Act Title I funding priorities and unmet service needs in Baltimore, Maryland. *Policy Stud. J.* 27(4):855–71
50. Kwasnicka D, Dombrowski SU, White M, Sniehotta F. 2016. Theoretical explanations for maintenance of behaviour change: a systematic review of behaviour theories. *Health Psychol. Rev.* 10:277–96
51. Lazer D, Pentland A, Adamic L, Aral S, Barabasi A-L, et al. 2009. Social science: computational social science. *Science* 323:721–23
52. Lerman K, Yan X, Wu X-Z. 2016. The "majority illusion" in social networks. *PLOS ONE* 11(2):e0147617
53. Lewis K, Kaufman J, Gonzalez M, Wimmer A, Christakis N. 2008. Tastes, ties, and time: a new social network dataset using Facebook.com. *Soc. Netw.* 30(4):330–42

54. Lomas J. 1991. Opinion leaders versus audit and feedback to implement practice guidelines: delivery after previous cesarean section. *JAMA* 265(17):2202–7
55. Luke DA, Harris JK. 2007. Network analysis in public health: history, methods, and applications. *Annu. Rev. Public Health* 28:69–93
56. Markham A, Buchanan E. 2012. *Ethical Decision-Making and Internet Research Recommendations from the AoIR Ethics Working Committee (Version 2.0)*. Chicago: Assoc. Internet Res.
57. McCormick TH, Lee H, Cesare N, Shojaie A, Spiro ES. 2015. Using Twitter for demographic and social science research: tools for data collection and processing. *Sociol. Methods Res.* **https://doi.org/10.1177/0049124115605339**
58. McNeill WH. 1989. *Plagues and Peoples*. New York: Anchor Books
59. Mercken L, Snijders TAB, Steglich C, Vertiainen E, De Vries H. 2010. Smoking-based selection and influence in gender-segregated friendship networks: a social network analysis of adolescent smoking: social network analysis of adolescent smoking. *Addiction* 105(7):1280–89
60. Moody J, Brynildsen WD, Osgood DW, Feinberg ME, Gest S. 2011. Popularity trajectories and substance use in early adolescence. *Soc. Netw.* 33(2):101–12
61. Morris M, Int. Union Sci. Study Popul., eds. 2004. *Network Epidemiology: A Handbook for Survey Design and Data Collection*. Oxford, UK/New York: Oxford Univ. Press
62. Mouttapa M, Valente T, Gallaher P, Rohrbach LA, Unger JB. 2004. Social network predictors of bullying and victimization. *Adolescence* 39(154):315–35
63. Natl. Cent. Chronic Dis. Prev. Health Promot. (US) Off. Smok. Health. 2012. *Preventing Tobacco Use Among Youth and Young Adults: A Report of the Surgeon General*. Atlanta: Cent. Dis. Control Prev.
64. Neaigus A, Friedman SR, Curtis R, Des Jarlais DC, Furst RT, et al. 1994. The relevance of drug injectors' social and risk networks for understanding and preventing HIV infection. *Soc. Sci. Med.* 38(1):67–78
65. Palinkas L, Spear SE, Mendon SJ, Villamar J, Valente TW, et al. 2015. Measuring sustainment of prevention programs and initiatives: a study protocol. *Implement. Sci.* 11:95
66. Pallotti F, Tubaro P, Casilli A, Valente TW. 2016. *"You see yourself like in a mirror": the effects of Internet-mediated personal networks on body image and eating disorders*. Work. Pap., Cent. Bus. Netw. Anal., Univ. Greenwich. **http://gala.gre.ac.uk/15685/1/2016_CBNA%232_Pallotti%20et%20al.pdf**
67. Poirier J, Cobb NK. 2012. Social influence as a driver of engagement in a web-based health intervention. *J. Med. Internet Res.* 14(1):e36
68. Provan KG, Milward HB. 1995. A preliminary theory of interorganizational network effectiveness: a comparative study of four community mental health systems. *Adm. Sci. Q.* 40(1):1–33
69. Richards P, Amara J, Ferme MC, Kamara P, Mokuwa E, et al. 2015. Social pathways for Ebola virus disease in rural Sierra Leone, and some implications for containment. *PLOS Negl. Trop. Dis.* 9(4):e0003567
70. Rogers EM, Kincaid DL. 1981. *Communication Networks: Toward a New Paradigm for Research*. New York: Free Press
71. Schölmerich VLN, Kawachi I. 2016. Translating the socio-ecological perspective into multilevel interventions: gaps between theory and practice. *Health Educ. Behav.* 43(1):17–20
72. Shalizi CR, Thomas AC. 2011. Homophily and contagion are generically confounded in observational social network studies. *Sociol. Methods Res.* 40(2):211–39
73. Shin H-S, Valente TW, Riggs NR, Huh J, Spruijt-Metz D, et al. 2014. The interaction of social networks and child obesity prevention program effects: the pathways trial. *Obesity* 22(6):1520–26
74. Siettos CI, Anastassopoulou C, Russo L, Grigoras C, Mylonakis E. 2016. Forecasting and control policy assessment for the Ebola virus disease (EVD) epidemic in Sierra Leone using small-world networked model simulations. *BMJ Open* 6(1):e008649
75. Snee H. 2008. *Web 2.0 as a Social Science Research Tool*. London: Br. Libr. ESRC Gov. Place. Scheme. **http://www.bl.uk/reshelp/bldept/socsci/socint/web2/web2.pdf**
76. Valente TW. 1995. *Network Models of the Diffusion of Innovations*. Cresskill, NJ: Hampton
77. Valente TW. 1996. Social network thresholds in the diffusion of innovations. *Soc. Netw.* 18(1):69–89
78. Valente TW. 2002. *Evaluating Health Promotion Programs*. New York: Oxford Univ. Press
79. Valente TW. 2010. *Social Networks and Health: Models, Methods, and Applications*. Oxford/New York: Oxford Univ. Press

80. Valente TW. 2012. Network interventions. *Science* 337:49–53
81. Valente TW. 2015. Social networks and health behavior. In *Health Behavior: Theory, Research, and Practice*, ed. K Glanz, BK Rimer, K Viswanath, pp. 205–22. New York: Wiley. 5th ed.
82. Valente TW, Coronges KA, Stevens GD, Cousineau MR. 2008. Collaboration and competition in a children's health initiative coalition: a network analysis. *Eval. Progr. Plan.* 31(4):392–402
83. Valente TW, Dougherty L, Stammer E. 2017. Response bias over time: interviewer learning and missing data in egocentric network surveys. Submitted
84. Valente TW, Dyal SR, Chu K-H, Wipfli H, Fujimoto K. 2015. Diffusion of innovations theory applied to global tobacco control treaty ratification. *Soc. Sci. Med.* 145:89–97
85. Valente TW, Fujimoto K, Palmer P, Tanjasiri SP. 2010. A network assessment of community-based participatory research: linking communities and universities to reduce cancer disparities. *Am. J. Public Health* 100(7):1319–25
86. Valente TW, Fujimoto K, Soto D, Ritt-Olson A, Unger JB. 2013. A comparison of peer influence measures as predictors of smoking among predominately Hispanic/Latino high school adolescents. *J. Adolesc. Health* 52(3):358–64
87. Valente TW, Palinkas LA, Czaja S, Chu K-H, Brown CH. 2015. Social network analysis for program implementation. *PLOS ONE* 10(6):e0131712
88. Valente TW, Rogers EM. 1995. The origins and development of the diffusion of innovations paradigm as an example of scientific growth. *Sci. Commun.* 16(3):242–73
89. Valente TW, Saba WP. 1998. Mass media and interpersonal influence in a reproductive health communication campaign in Bolivia. *Commun. Res.* 25(1):96–124
90. Valente TW, Watkins SC, Jato MN, Van Der Straten A, Tsitsol L-PM. 1997. Social network associations with contraceptive use among Cameroonian women in voluntary associations. *Soc. Sci. Med.* 45(5):677–87
91. Vanhems P, Barrat A, Cattuto C, Pinton J-F, Khanafer N, et al. 2013. Estimating potential infection transmission routes in hospital wards using wearable proximity sensors. *PLOS ONE* 8(9):e73970
92. Wesolowski A, Buckee CO, Bengtsson L, Wetter E, Lu X, Tatem AJ. 2014. Commentary: Containing the Ebola outbreak—the potential and challenge of mobile network data. *PLOS Curr*. Vol. 6. **https://doi.org/10.1371/currents.outbreaks.0177e7fcf52217b8b634376e2f3efc5e**
93. West E, Barron DN, Dowsett J, Newton JN. 1999. Hierarchies and cliques in the social networks of health care professionals: implications for the design of dissemination strategies. *Soc. Sci. Med.* 48(5):633–46
94. Work. Party Internet-Mediat. Res. 2013. *Ethics Guidelines for Internet-Mediated Research*. Leicester, UK: Br. Psychol. Soc.
95. Yeung K-T, Martin JL. 2003. The looking glass self: an empirical test and elaboration. *Soc. Forces* 81(3):843–79
96. Zappa P, Lomi A. 2015. The analysis of multilevel networks in organizations: models and empirical tests. *Organ. Res. Methods.* 18(3):542–69
97. Zimmer M. 2010. "But the data is already public": on the ethics of research in Facebook. *Ethics Inf. Technol.* 12(4):313–25