

# **HHS Public Access**

Author manuscript *HERD*. Author manuscript; available in PMC 2022 October 01.

Published in final edited form as:

HERD. 2021 October ; 14(4): 18-34. doi:10.1177/19375867211013772.

# Best Practices for Modeling Egocentric Social Network Data and Health Outcomes

Jacqueline M. Burgette, DMD, PhD<sup>a</sup>, Jacquelin Rankine, MD, MS<sup>b</sup>, Alison J. Culyba, MD, PhD, MPH<sup>c</sup>, Kar-Hai Chu, PhD<sup>d</sup>, Kathleen M. Carley, PhD<sup>e</sup>

<sup>a</sup>Assistant Professor, Departments of Dental Public Health and Pediatric Dentistry, University of Pittsburgh School of Dental Medicine, 3501 Terrace Street, Pittsburgh, PA 15261, USA

<sup>b</sup>Fellow, Division of Adolescent and Young Adult Medicine, Department of Pediatrics, University of Pittsburgh School of Medicine, 120 Lytton Avenue, Pittsburgh, Pennsylvania 15213, USA

<sup>c</sup>Assistant Professor, Division of Adolescent and Young Adult Medicine, Department of Pediatrics, University of Pittsburgh School of Medicine, 120 Lytton Avenue, Pittsburgh, Pennsylvania 15213, USA

<sup>d</sup>Associate Professor, Behavioral and Community Health Sciences, University of Pittsburgh Graduate School of Public Health, 130 DeSoto Street, Pittsburgh, PA 15261, USA

<sup>e</sup>Professor, School of Computer Science, Center for Computational Analysis of Social and Organizational Systems, Carnegie Mellon University, Wean Hall 5130, 5000 Forbes Ave, Pittsburgh, PA 15213, USA

# Keywords

Health; Models; Statistical; Social Networking; Ego; Humans; Social Support

# Introduction

Social Network Analysis (SNA) is a scientific area focused on the study of relations. Using SNA, we can model and analyze these relations to understand how they influence attitudes, experience, or behavior—and subsequently health. The growing interest in SNA in health research is not surprising, as many health processes are inherently relational. Consider the spread of a contagious disease through close social or environmental contacts, the sharing of health information with friends or family, or the likelihood that a teenager adopts the risky health behaviors of peers. SNA provides unique approaches to study the impact of social relationships, and the environmental contexts in which social interactions occur, on these and other health processes in a way not afforded by traditional methods.

Previous research has described associations between features of social and built environments—such as population density, diversity of land use, and urban design—and

Conflict of Interest: The authors have no conflicts of interest to report.

Corresponding Author: Jacqueline M. Burgette DMD, PhD, 3501 Terrace Street, Pittsburgh, PA 15261, Phone: (412) 648-8504, Fax: (412) 383-8662, jacqueline@pitt.edu.

social network structures (Cabrera and Najarian, 2015; Boessen et al., 2018). The combined impact of the environment and social networks has been shown to affect health behaviors, such as physical inactivity and gang violence (Papachristos et al., 2013; Josey and Moore, 2018). These studies support Carpiano's resource-based theory of social capital for health in which structural characteristics of a neighborhood and its surrounding areas, or "structural antecedents," lay the groundwork for the actual or potential resources rooted in social networks (Carpiano, 2006). From these social networks, positive or negative resources can be received to impact health and well-being (Carpiano, 2006).

Understanding of network-based data is of particular importance in research on the health environment. There is increasing focus on the dynamic interplay of interpersonal, social, and environmental influences on health. In addition to modeling relational health processes, SNA lends itself to exploration of these multilevel influences on the determinants and spread of health and disease. In SNA, health behaviors and outcomes may be modeled as emergent properties of the larger network system to gain a unique understanding of social influences on health and disease. SNA approaches have been applied to several domains of health research including epidemiology and public health (Smith & Christakis, 2008), health organizational behavior (Merrill et al., 2007, 2008, 2010, 2012), healthcare quality and safety (Brewer et al., 2018; Effken et al., 2011), and health services delivery (Provan et al., 2010).

Traditional research approaches focused on social influences on health attempt to predict health outcomes based on attribute data, i.e., characteristics or behaviors of individuals or groups within one's social sphere. Established survey instruments on social connections assess only limited relational data such as an individual's number, type, or quality of relations overall (Chang et al., 2018; Cohen et al., 1997). Network-based data tells us something inherently different than general measures of social support or social capital obtained through non-network attribute or relational data. SNA instead focuses on the structure of an individual's web of relationships, in addition to specific features of each relationship, and how they can affect health outcomes.

Applying SNA methods to research on the health environment requires understanding of network-based terminology, and strengths and limitations of analytic approaches. There are two primary types of social networks. *Sociocentric*, or *whole networks*, include all members of a bounded network (e.g., all students in a school) and the exhaustive web of relations between them. *Egocentric*, or *personal networks*, take the perspective of an individual actor —termed *ego*—who identifies their own connections—termed *alters*—and the relationships between these alters. It is often more feasible in public or population health research to collect egocentric network data using traditional epidemiologic sampling approaches. Therefore, the focus of this manuscript will be on the analysis of egocentric networks.

There is a small but growing collection of studies utilizing egocentric network methods in health research. Egocentric network methods have been applied to sexual and HIV transmission risk behaviors (Felsher & Koku, 2018; Kennedy et al., 2012; Latkin et al., 2011; Li et al., 2011; Morris et al., 2010), violence perpetration and victimization (Bond & Bushman, 2017; Haynie et al., 2006), substance use and smoking behaviors (Huang et al.,

2014; Li et al., 2012; Takagi et al., 2020; Valente & Vlahov, 2001), depression and mental health (Child & Lawton, 2020; Lam et al., 2017), contraceptive use (Valente & Saba, 1988), obesity (Christakis & Fowler, 2009; Hruschka et al., 2011), health promoting behavior (Flatt et al., 2012; Kennedy-Hendricks et al., 2015; O'Malley et al., 2012), dental health (Pullen et al., 2018), healthcare utilization (Adams et al., 2015; Cheong et al., 2013), and healthcare coordination (McGhee Hasrick et al., 2018).

While there are some common measures frequently used in modeling the impact of egocentric social networks on health outcomes, there is no standard approach to egocentric network modeling in health research, which impacts rigor and reproducibility. This guide provides an overview of social network concepts and terminology and outlines best practices to standardize and strengthen future research on the health environment utilizing egocentric network methods.

# **Concepts and Terminology**

For health environment researchers aiming to incorporate Social Network Analysis (SNA) approaches into their research, we present a limited review of guiding SNA principles and a glossary of network-based terminology (see Table 1). For a comprehensive introduction to social network theory and analysis, interested readers may wish to consult the following sources (Altman et al., 2018; Carrasco et al., 2008; Hawe et al., 2004; Hogan et al., 2007; Luke & Harris, 2007; Scott, 2017; Valente, 2010; Wasserman & Faust, 1994; Wellman & Frank, 2001).

#### **Representing Social Network Data**

In SNA, data is collected to identify *actors* and the *relational ties* between them. Actors and their relational ties exist within *networks*. Actors are individual entities such as people, groups, or organizations. An *ego* is the focal actor of interest, from whose perspective the network is conceptualized, and *alters* are all of the other actors within the network. Relational ties are the connections between actors (i.e. ego-alter and alter-alter connections). Relational ties may carry characteristics of the relationship such as the strength (e.g., emotional closeness, level of trust, frequency of communication), type (e.g., friends, coworkers, kin), direction (e.g., mutual, unreciprocated), or function (e.g., health information sharing, financial support). Actors and their relational ties exist within *networks*. Once relational data is collected within a network, it can be recorded in a square actor by actor *matrix* (see Figure 1a). Using SNA software, a matrix of relational data may be represented visually as *graphs* for further exploration and analysis (see Figure 1b). Actors and relational ties are displayed in network graphs as *nodes* and *links* respectively.

**Units of Measurement within Social Network Analysis**—Social network measures may be determined for expanding units of the network, from an individual actor to the network as a whole. *Individual* measures identify the position or role of a single actor within a network. Individual measures of position (e.g., degree, betweenness) or role (e.g., isolate, bridge) can help to understand the likelihood that an actor receives or transmits information, behavior, or disease. *Network* measures describe the structure of all actors and relational ties within the network as a whole (e.g., density, centralization) to make inferences about the

experiences of network members. Depending on the research question of interest, network measures can be calculated from the perspective of the individual actor and/or the network.

**Unique Concepts for SNA in Health Environment Research**—Social Network Analysis (SNA) allows for unique approaches to study the determinants and spread of health and disease within populations and environments. In health research, social networks are vehicles through which people can pass ideas (e.g., health attitudes, norms, information), support (e.g., social or financial support, collaboration), resources, health behaviors, health services, and the disease itself. Generally, relational ties, also known as network links, can be thought of as channels for the transfer or flow of elements (e.g. ideas, support, resources, behaviors, pathogens) between actors (Wasserman & Faust, 1994). Three primary categories of network properties can help to understand these transfers as they relate to health and disease: access, brokerage, and social influence. An understanding of these network properties and the associated health processes is needed to guide measure selection in SNA as detailed below.

# Best Practices for Modeling Egonetworks and Health Outcomes

#### Step 1: Model Selection

Regression is a broad collection of statistical models that allow the researcher to estimate how a response variable changes as a function of other variables, and to assess the statistical significance of such differences. Regression-based approaches are frequently used to examine the relationship between a social network exposure and a health outcome using egocentric network data. Consistent with statistical modeling in general, the type of regression depends on the classification of the main outcome variable being used in the analysis. For example, if the main outcome variable is binary, then a logit or probit model may be appropriate. If the main outcome variable is a count, then a Poisson or negative binomial model may be appropriate. Researchers should be mindful of study design (e.g clustered sampling) and select regression models suited to the data structure.

Network models such as exponential random graph models (ERGMs) and multiple regression quadratic assignment procedures (MRQAP) may be better suited for analyzing sociocentric networks. There is recent literature on the use of ERGMs to analyze egocentric network data using advanced statistics and programming (Box-Steffensmeier et al., 2018; Krivitsky & Morris, 2017). For the purposes of this manuscript, we seek to assist health researchers who may not have a rich background in network science and will therefore focus on regression-based approaches.

#### Step 2: Social Network Exposure Variable(s) and Selection Considerations

Network properties can be broadly organized into three overarching categories pertinent to health outcomes: access, brokerage, and social influence. We suggest health researchers select access, brokerage, and/or social influence as the primary exposure(s) of interest based on the primary research question and postulated links between social network properties and the health outcome of interest. We present an overview of metrics for assessing these

network properties and propose best practices for social network measure selection within each category.

**Access.**—Measures of *access* broadly serve to predict how likely an actor is to receive or transmit any element, including health information, resources, and services. Access measures are typically individual measures of *centrality*, the prominence of an actor within a network as determined by how extensively connected that actor is to other network members (Table 2). Centrality measures of access include *degree* and *closeness*. Degree is the total number of actors directly connected to the ego. Degree captures the popularity or activity of an ego. Actors with high degree are directly connected to more other actors, and therefore have greater access to information, support, or resources. Closeness is the inverse sum of the distance in links connecting the ego to all other actors in the network. Alters with high closeness can rapidly interact with all other actors and are less likely to rely on intermediary actors to quickly access their network.

To summarize, actors with more ties to other actors, or shorter paths connecting them to other actors, have greater access. Bond & Bushman (2017) applied measures of access to the study of violence among adolescents to test the theory that violent behaviors are transmitted along ties in a social network—much like a contagious disease. Adolescents were more likely to display violent behaviors when they had easier access to observed violent behaviors through direct ties (i.e., greater degree) or shorter path connections (i.e., greater closeness) to violent peers, with measurable but waning associations up to four degrees of separation.

**Brokerage.**—Measures of *brokerage* identify actors that are likely to be influential due to their positions along key paths of transmission, acting as gatekeepers or conduits of information or disease. Measures of brokerage include *betweenness* and the number of *structural holes*. Betweenness is the extent to which the ego connects pairs of other actors by falling on the shortest path between these actors. Egos with high betweenness may be thought of as intermediaries that control, or broker, the flow of elements between the actors that they connect. Because of their position on many shortest paths, egos with high betweenness will also be frequently exposed to any elements transmitted along these paths.

Structural holes are a special form of brokerage where an ego is a member of two groups (e.g., work and family groups) that are not connected except for through the ego. An egocentric network with structural holes will allow the ego greater access to novel information or resources from the unconnected groups. In addition, the ego may have more freedom to act autonomously without the check of a highly interconnected network. For example, an ego can act one way with work colleagues and another way with family without the awareness of the other group.

To summarize, an actor in a high brokerage position may act as a gatekeeper along paths between many other actors, sit in a highly trafficked position, or bridge two or more otherwise unconnected components. Measures of brokerage have been used to study the network-based transmission of infectious diseases including HIV (Christley et al., 2005; Gyarmathy et al., 2014). High brokerage actors were more likely to contract HIV themselves due to greater exposure as a result of their high traffic positions (i.e., high betweenness).

High brokerage actors were also key to community spread of HIV as a result of their bridging positions between otherwise unconnected components (i.e., spanning of structural holes).

**Social influence.**—Social influence is a process by which an ego is impacted by their alters, who are also connected to and influenced by others. Measures of social influence can estimate the extent to which an ego's beliefs, behaviors and/or disease states are a function of the ego's connection to other individuals. Social influence measures can be calculated for the network as a whole and include measures of *cohesion*, the interconnectedness of actors within a network. The primary cohesion measure, *density*, captures whether an egocentric network is close-knit. Density is the number of links present in a network divided by all possible links which could exist. Network size and density are often negatively correlated; as size increases, density decreases.

Social influence can also be approximated through network-level measures of link strength and direction including *average link strength*, *reciprocity*, and *transitivity*. Average link strength is the network-level mean of link strength, a value measure assigned to a link based on the intensity, duration, or frequency of contact. Based on the research question of interest, various prompts can be used to obtain ego-reported ratings of link strength and direction for each relationship during data collection. Links may be unidirectional from alter to ego (e.g., Does this alter give you health advice?), unidirectional from ego to alter (e.g., Do you give this alter health advice?), or bidirectional. Reciprocity is the percentage of links in a network that are bidirectional. Strong or bidirectional links can be assumed to exert greater forces of social influence. Determining link direction also allows for identification of triads and measurement of transitivity, which is reviewed in other literature (Faust, 2006; Snijders, 2011).

Finally, social influence can be captured using *network similarity* measures to account for the extent to which an ego is similar to their alters. Network similarity can be operationalized as the proportion of the ego's alters that share a certain characteristic. If an alter attribute is binary, for example any lifetime tobacco use, it is useful to calculate the proportion of alters within an ego network with that attribute. If an alter attribute is a continuous measure, for example body mass index, it is useful to calculate a mean and standard deviation if normally distributed or a median and interquartile range if not normally distributed.

To summarize, actors within close-knit networks or actors connected by strong or reciprocal links will experience greater social influence. Actors within networks where many other actors share similar characteristics will also experience greater social influence. Social influence may be beneficial if the social force transmitted is social support or social capital. Social influence may also be detrimental to health. Situations of high social influence make it more challenging for an actor to adopt or maintain behaviors opposite of its network—the network equivalent of peer pressure. Lam et al., (2017) evaluated the role of measures of social influence. Women in close-knit networks (i.e., high density) of social support had decreased odds of depression, while women within highly similar networks (i.e., high

network similarity) with regards to unhealthy behaviors (e.g., smoking) had increased odds of depression.

#### Step 3: Covariate Selection: Sociodemographic and Health Characteristics

Similar to traditional regression modeling, researchers can select covariates based on the literature. For example, having dental insurance is associated with increased personal use of dental services. Therefore, models that describe dental utilization often control for dental insurance. In the absence of data, theory and model diagnostics should be used to guide the selection of covariates that are relevant to the health outcome. These variables may be on the individual-, family-, community- and country- level.

#### Step 4: Covariate Selection: Social Network Characteristics

In addition to serving as the main exposure variable of interest, network measures can also be covariates. When examining the impact of social networks on health outcomes, such as the impact of brokerage on acquiring a disease, the researcher will need to use both knowledge from previous literature and theory related to the health outcome in order to determine which social network measures should to be included as exposures or covariates. Failure to include relevant network measures as covariates can lead to model misspecification. Researchers can run the regression model with and without network variables to test whether the inclusion of network variables impacts the model. The field of health and social networks has progressed from early studies that only controlled for individual-level covariates to more recent advances in social network modeling that highlight the importance of also including social network covariates (Merrill et al., 2015). We recommend the best practice of including social network covariates into social network models moving forward.

Many measures of access, brokerage, and social influence are impacted by network size. As a general rule, all other network measures depend on network size mathematically, so this measure affects all other network measures. It is therefore best practice to include network size as a covariate in analytic models.

There should be a rationale for including or excluding each network covariate in the analytic model. In addition to network size, there are three main categories of network variables to consider as model covariates: centrality, cohesion, and similarity measures. The inclusion of each category of social network covariates should be based on whether the category is perceived as a driver of the health outcome. For example, models of health outcomes that are dependent on contact, such as contagious diseases, may choose to include network size and cohesion measures. Researchers should use theory to guide which variables are relevant to a particular health outcome in the absence of previous social network literature.

When choosing model covariates, the network covariates must not only be consistent with previous literature and theory, but also be statistically rigorous. For example, controlling for both total degree and betweenness may not be statistically rigorous due to collinearity. Centrality and cohesion measures may be highly correlated, so we recommend the best practice of checking for collinearity. If a researcher prefers not to select one centrality

measure, more than one centrality measure can be combined, for example, by using factor analysis.

#### **Step 5: Analytic Considerations**

There are a number of analytic issues to consider depending on the sampling approach and specifics of the network data. We will outline a few common topics to consider during data analysis. The selection of network measures may be influenced by both the number of egos in the study and the size of each ego network. For example, if many ego networks are small (i.e., each network having fewer than three alters) we advise against using betweenness or density because it may be difficult to discriminate between ego networks.

One of the many assumptions underlying regression models is that participants are independent (Rothman et al., 2008). Non-independence can become increasingly complex in ego network analysis. Network data can involve egos that have relationships with the same alters (Smith, 2012), or the egos themselves could be linked (Illenberger et al., 2011). If independent egos A and B both list person C as an alter, then any network calculation for A will no longer be independent from B. Additionally, the characteristics of the alters, or alter and ego, may not be independent (Wellman, 2007).

As with non-network studies, non-independence can also arise from clustered sampling. Specific regression models (e.g. linear mixed models and generalized estimating equations) can be employed to model egocentric network exposures and health outcomes in clustered studies. These models account for the inherent similarities between participants within a given cluster (e.g. school) and can be applied similarly to how they are used with non-network-based exposures.

Researchers continue to work on how to avoid or counter the effects of such lacks of independence in regression-based ego network analysis. Generally, the strengths, availability, and familiarity with regression-based approaches outweigh this limitation, and favor regression-based modeling. We recommend noting and discussing this as a limitation when relevant.

Finally, this manuscript is focused on network data at one time point. Dynamic panel network data requires advanced considerations about metric stability over time (Benham-Hutchins et al., 2018) beyond the scope of these recommendations.

# Examples

To illustrate the best practices described in the manuscript, we applied the five-step process to examine whether social influence is associated with child dental utilization (see Figure 2 and Table 3) (Pullen et al., 2018; Fisher-Owens et al., 2007). We chose two social network exposure variables of social influence to illustrate how the same research question can be approached using different network exposure variables. In the first example, we examined whether more network similarity on dental utilization among a mother's alters is associated with improved dental utilization for her child. In the second example, we examined whether less network density in a mother's social network is associated with

improved dental utilization for her child. Geography and the built environment can affect whether the mother's alters have access to dental care and the density of the mother's social network. In these simplified examples, we performed adjusted logistic regressions controlling for 4 individual- and family-level covariates and 4 network covariates.

# Discussion

Health researchers are increasingly recognizing the impact of social relationships on health outcomes, including the combined impact of the environment and social networks on health. Over the past two decades, there has been a greater focus on applying Social Network Analysis (SNA) in health research to better understand how individuals' health outcomes can be affected by their social networks, to develop network-informed interventions, and to open new avenues of research based on network modeling. We developed this guide to foster both conceptual understanding as well as practical application for SNA in health environment research. It is meant to both provide experienced researchers with standardized methodologies that can be translatable and reproducible, as well as encourage new researchers who are unfamiliar with SNA to develop high quality analyses in health and the environment. By outlining the steps from model selection, to selecting exposure variables and covariates, through analytic considerations, we wanted this process to feel familiar to traditional health researchers while simultaneously offering the concepts, metrics, and analyses that SNA affords.

It is imperative that high quality modeling adheres to best practices that produce generalizable and reproducible results. Researchers must also have a clear understanding of how and why social network methods can expand beyond classic social influence factors (e.g., behaviors or perceived support), and utilize social network data that emphasizes the importance of relationship structures and how they can affect the individual's unique position in the network. We believe this guide and example highlight that delineation and clarify how SNA can provide a different insight into health research.

Social network research has impacted both health services and policies. SNA can also be used to inform interventions that are methodologically rigorous and result in improved health outcomes (Valente, 2012). For example, peer educators were identified to train their social network of drug or sex partners on reducing HIV risk behaviors (Latkin et al., 2009) and student-nominated opinion leaders were able to successfully lead tobacco prevention programming in high schools (Valente et al., 2003). These, and other studies, apply concepts that we have discussed (e.g., social influence) and leverage social networks to positively impact health outcomes. There remains significant room to advance SNA informed public health interventions by promoting standardized SNA approaches which conform to best practices for modeling social network exposures and health outcomes as outlined here (Shelton et al., 2009). For example, social network analysis can be used to optimize the design of health organizations or health facilities to improve health outcomes and patient safety (Effken et al., 2011; Benham-Hutchins et al., 2018; Brewer et al., 2018). Many of the concepts and theories in this field of work are still evolving; we hope this review is a timely introduction to the application of SNA in health research, particularly related to the health environment.

# Funding Sources:

This work was supported by the Robert Wood Johnson Foundation Harold Amos Medical Faculty Development Program and the National Institute of Health, Eunice Kennedy Shriver National Institute of Child Health and Human Development (K23 HD098277–01).

# References

- Adams A, Nababan H, & Hanifi S (2015). Building Social Networks for Maternal and Newborn Health in Poor Urban Settlements: A Cross-Sectional Study in Bangladesh. PLOS ONE, 10(4), e0123817. doi: 10.1371/journal.pone.0123817 [PubMed: 25910191]
- Altman N, Carley KM, & Reminga J (2018). ORA User's Guide 2018. Technical Report CMU-ISR-18–103 Pittsburgh, PA: Carnegie Mellon University, School of Computer Science, Institute for Software Research. http://www.casos.cs.cmu.edu/projects/ora/CMU-ISR-18-103.pdf
- Benham-Hutchins M, Carley K, Brewer B, Effken J, & Reminga J (2018). Nursing Unit Communication During a US Public Health Emergency: Natural Experiment. JMIR Nursing, 1(1), e11425. doi: 10.2196/11425 [PubMed: 34345768]
- Brewer B, Carley K, Benham-Hutchins M, Effken J, Reminga J, & Kowalchuck M (2018). Relationship of Staff Information Sharing and Advice Networks to Patient Safety Outcomes. JONA: The Journal Of Nursing Administration, 48(9), 437–444. doi: 10.1097/nna.00000000000646
- Bond R, & Bushman B (2017). The Contagious Spread of Violence Among US Adolescents Through Social Networks. American Journal Of Public Health, 107(2), 288–294. doi: 10.2105/ ajph.2016.303550 [PubMed: 27997233]
- Boessen A, Hipp JR, Butts CT, Nagle NN & Smith EJ (2018). The Built Environment, Spatial Scale, and Social Networks: Do Land Uses Matter for Personal Network Structure? Environment and Planning B: Urban Analytics and City Science, 45(3) 400–416. doi: 10.1177/2399808317690158
- Box-Steffensmeier J, Campbell B, Christenson D, & Navabi Z (2018). Role analysis using the ego-ERGM: A look at environmental interest group coalitions. Social Networks, 52, 213–227. doi: 10.1016/j.socnet.2017.08.004
- Cabrera JF, & Najarian JC (2015). How the Built Environment Shapes Spatial Bridging Ties and Social Capital. Environment and Behavior, 47(3) 239–267. doi: 10.1177/0013916513500275
- Carpiano RM (2006). Toward a Neighborhood Resource-Based Theory of Social Capital for Health: Can Bourdieu and Sociology Help? Social Science & Medicine, 62(1):165–75. doi: 10.1016/ j.socscimed.2005.05.020 [PubMed: 15992978]
- Carrasco J, Hogan B, Wellman B, & Miller E (2008). Collecting Social Network Data to Study Social Activity-Travel Behavior: An Egocentric Approach. Environment And Planning B: Planning And Design, 35(6), 961–980. doi: 10.1068/b3317t
- Chang Q, Sha F, Chan C, & Yip P (2018). Validation of an abbreviated version of the Lubben Social Network Scale ("LSNS-6") and its associations with suicidality among older adults in China. PLOS ONE, 13(8), e0201612. doi: 10.1371/journal.pone.0201612 [PubMed: 30071067]
- Cheong L, Armour C, & Bosnic-Anticevich S (2013). Primary health care teams and the patient perspective: A social network analysis. Research In Social And Administrative Pharmacy, 9(6), 741–757. doi: 10.1016/j.sapharm.2012.12.003 [PubMed: 23395532]
- Child S, & Lawton L (2020). Personal networks and associations with psychological distress among young and older adults. Social Science & Medicine, 246, 112714. doi: 10.1016/ j.socscimed.2019.112714 [PubMed: 31864967]
- Christakis N, & Fowler J (2007). The Spread of Obesity in a Large Social Network over 32 Years. New England Journal Of Medicine, 357(4), 370–379. doi: 10.1056/nejmsa066082
- Christley R, Pinchbeck G, Bowers R, Clancy D, French N, Bennett R, & Turner J (2005). Infection in Social Networks: Using Network Analysis to Identify High-Risk Individuals. American Journal Of Epidemiology, 162(10), 1024–1031. doi: 10.1093/aje/kwi308 [PubMed: 16177140]
- Cohen S (1997). Social ties and susceptibility to the common cold. JAMA: The Journal Of The American Medical Association, 277(24), 1940–1944. doi: 10.1001/jama.277.24.1940 [PubMed: 9200634]

- Effken J, Carley K, Gephart S, Verran J, Bianchi D, Reminga J, & Brewer B (2011). Using ORA to explore the relationship of nursing unit communication to patient safety and quality outcomes. International Journal Of Medical Informatics, 80(7), 507–517. doi: 10.1016/j.ijmedinf.2011.03.015 [PubMed: 21536492]
- Faust K (2006). Comparing social networks: Size, density and local structure. Metodološki Zvezki, Advances in Methodology and Statistics, 3(2):185–216.
- Felsher M, & Koku E (2018). Explaining HIV Risk Multiplexity: A Social Network Analysis. AIDS And Behavior, 22(11), 3500–3507. doi: 10.1007/s10461-018-2120-7 [PubMed: 29680933]
- Fisher-Owens S, Gansky S, Platt L, Weintraub J, Soobader M, Bramlett M, & Newacheck P (2007). Influences on Children's Oral Health: A Conceptual Model. PEDIATRICS, 120(3), e510–e520. doi: 10.1542/peds.2006-3084 [PubMed: 17766495]
- Flatt J, Agimi Y, & Albert S (2012). Homophily and Health Behavior in Social Networks of Older Adults. Family & Community Health, 35(4), 312–321. doi: 10.1097/fch.0b013e3182666650 [PubMed: 22929377]
- Gyarmathy V, Caplinskiene I, Caplinskas S, & Latkin C (2014). Social Network Structure and HIV Infection Among Injecting Drug Users in Lithuania: Gatekeepers as Bridges of Infection. AIDS And Behavior, 18(3), 505–510. doi: 10.1007/s10461-014-0702-6 [PubMed: 24469223]
- Hawe P (2004). A glossary of terms for navigating the field of social network analysis. Journal Of Epidemiology & Community Health, 58(12), 971–975. doi: 10.1136/jech.2003.014530 [PubMed: 15547054]
- Haynie D, Silver E, & Teasdale B (2006). Neighborhood Characteristics, Peer Networks, and Adolescent Violence. Journal Of Quantitative Criminology, 22(2), 147–169. doi: 10.1007/s10940-006-9006-y
- Hogan B, Carrasco J, & Wellman B (2007). Visualizing Personal Networks: Working with Participantaided Sociograms. Field Methods, 19(2), 116–144. doi: 10.1177/1525822×06298589
- Hruschka D, Brewis A, Wutich A, & Morin B (2011). Shared Norms and Their Explanation for the Social Clustering of Obesity. American Journal Of Public Health, 101(S1), S295–S300. doi: 10.2105/ajph.2010.300053 [PubMed: 21555656]
- Huang G, Unger J, Soto D, Fujimoto K, Pentz M, Jordan-Marsh M, & Valente T (2014). Peer Influences: The Impact of Online and Offline Friendship Networks on Adolescent Smoking and Alcohol Use. Journal Of Adolescent Health, 54(5), 508–514. doi: 10.1016/ j.jadohealth.2013.07.001
- Illenberger J, Kowald M, Axhausen K, & Nagel K (2011). Insights into a spatially embedded social network from a large-scale snowball sample. The European Physical Journal B, 84(4), 549–561. doi: 10.1140/epjb/e2011-10872-0
- Josey MJ, & Moore S (2018). The Influence of Social Networks and The Built Environment on Physical Inactivity: A Longitudinal Study of Urban-Dwelling Adults. Health Place, 54:62–68. doi: 10.1016/j.healthplace.2018.08.016. [PubMed: 30245244]
- Kennedy D, Tucker J, Green H, Golinelli D, & Ewing B (2012). Unprotected Sex of Homeless Youth: Results from a Multilevel Dyadic Analysis of Individual, Social Network, and Relationship Factors. AIDS And Behavior, 16(7), 2015–2032. doi: 10.1007/s10461-012-0195-0 [PubMed: 22610421]
- Kennedy-Hendricks A, Schwartz H, Thornton R, Griffin B, Green H, & Kennedy D et al. (2015). Intergenerational Social Networks and Health Behaviors Among Children Living in Public Housing. American Journal Of Public Health, 105(11), 2291–2297. doi: 10.2105/ ajph.2015.302663 [PubMed: 26378821]
- Krivitsky P, & Morris M (2017). Inference for social network models from egocentrically sampled data, with application to understanding persistent racial disparities in HIV prevalence in the US. The Annals Of Applied Statistics, 11(1), 427–455. doi: 10.1214/16-aoas1010 [PubMed: 29276550]
- Lam J, Lu B, Doogan N, Thomson T, Ferketich A, Paskett E, & Wewers M (2017). Depression, smoking, and ego-centric social network characteristics in Ohio Appalachian women. Journal Of Rural Mental Health, 41(1), 30–41. doi: 10.1037/rmh0000054 [PubMed: 29081878]

- Latkin C, Donnell D, Metzger D, Sherman S, Aramrattna A, & Davis-Vogel A et al. (2009). The efficacy of a network intervention to reduce HIV risk behaviors among drug users and risk partners in Chiang Mai, Thailand and Philadelphia, USA. Social Science & Medicine, 68(4), 740–748. doi: 10.1016/j.socscimed.2008.11.019 [PubMed: 19070413]
- Latkin C, Yang C, Tobin K, Penniman T, Patterson J, & Spikes P (2011). Differences in the Social Networks of African American Men Who Have Sex With Men Only and Those Who Have Sex With Men and Women. American Journal Of Public Health, 101(10), e18–e23. doi: 10.2105/ ajph.2011.300281 [PubMed: 21852650]
- Li J, Liu H, Li J, Luo J, Koram N, & Detels R (2011). Sexual transmissibility of HIV among opiate users with concurrent sexual partnerships: an egocentric network study in Yunnan, China. Addiction, 106(10), 1780–1787. doi: 10.1111/j.1360-0443.2011.03459.x [PubMed: 21457169]
- Li J, Weeks M, Borgatti S, Clair S, & Dickson-Gomez J (2012). A Social Network Approach to Demonstrate the Diffusion and Change Process of Intervention From Peer Health Advocates to the Drug Using Community. Substance Use & Misuse, 47(5), 474–490. doi: 10.3109/10826084.2012.644097 [PubMed: 22428816]
- Luke D, & Harris J (2007). Network Analysis in Public Health: History, Methods, and Applications. Annual Review Of Public Health, 28(1), 69–93. doi: 10.1146/ annurev.publhealth.28.021406.144132
- McGhee Hassrick E, Shattuck P, & Carley K (2018). Network Measures of Collaborative Support for Young Adults With Autism. Pediatrics, 141(Supplement 4), S287–S292. doi: 10.1542/ peds.2016-4300e [PubMed: 29610409]
- Merrill J, Bakken S, Rockoff M, Gebbie K, & Carley K (2007). Description of a method to support public health information management: Organizational network analysis. Journal Of Biomedical Informatics, 40(4), 422–428. doi: 10.1016/j.jbi.2006.09.004 [PubMed: 17098480]
- Merrill J, Caldwell M, Rockoff M, Gebbie K, Carley K, & Bakken S (2008). Findings from an Organizational Network Analysis to Support Local Public Health Management. Journal Of Urban Health, 85(4), 572–584. doi: 10.1007/s11524-008-9277-8 [PubMed: 18481183]
- Merrill J, Keeling J, & Carley K (2010). A Comparative Study of 11 Local Health Department Organizational Networks. Journal Of Public Health Management And Practice, 16(6), 564–576. doi: 10.1097/phh.0b013e3181e31cee [PubMed: 20445462]
- Merrill J, Orr M, Jeon C, Wilson R, Storrick J, & Carley K (2012). Topology of Local Health Officials' Advice Networks. Journal Of Public Health Management And Practice, 18(6), 602–608. doi: 10.1097/phh.0b013e31825d20ac [PubMed: 23023286]
- Morris M, Epstein H, & Wawer M (2010). Timing Is Everything: International Variations in Historical Sexual Partnership Concurrency and HIV Prevalence. Plos ONE, 5(11), e14092. doi: 10.1371/ journal.pone.0014092 [PubMed: 21124829]
- O'Malley A, Arbesman S, Steiger D, Fowler J, & Christakis N (2012). Egocentric Social Network Structure, Health, and Pro-Social Behaviors in a National Panel Study of Americans. Plos ONE, 7(5), e36250. doi: 10.1371/journal.pone.0036250 [PubMed: 22615760]
- Papachristos AV, Hureau DM, & Braga AA (2013). The Corner and the Crew: The Influence of Geography and Social Networks on Gang Violence. American Sociological Review, 78(3) 417– 447. doi: 10.1177/0003122413486800
- Provan K, Leischow S, Keagy J, & Nodora J (2010). Research collaboration in the discovery, development, and delivery networks of a statewide cancer coalition. Evaluation And Program Planning, 33(4), 349–355. doi: 10.1016/j.evalprogplan.2009.12.005 [PubMed: 20061027]
- Pullen E, Perry B, & Maupome G (2018). "Does this Look Infected to You?" Social Network Predictors of Dental Help-Seeking Among Mexican Immigrants. Journal Of Immigrant And Minority Health, 20(2), 399–409. doi: 10.1007/s10903-017-0572-x [PubMed: 28361383]
- Rothman KJ, Greenland S, & Lash TL (2008). Modern Epidemiology (3rd ed.). Philadelphia, PA: Lippincott, Williams and Wilkins.
- Scott J (2017) Social Network Analysis (4th ed.). Thousand Oaks, CA: Sage Publications, Inc.
- Sheehan B, Carley K, Stetson P, & Merrill J (2015). Transition Networks in a Cohort of Patients with Congestive Heart Failure. Applied Clinical Informatics, 06(03), 548–564. doi: 10.4338/ aci-2015-02-ra-0021

- Shelton R, Lee M, Brotzman L, Crookes D, Jandorf L, Erwin D, & Gage-Bouchard E (2019). Use of social network analysis in the development, dissemination, implementation, and sustainability of health behavior interventions for adults: A systematic review. Social Science & Medicine, 220, 81–101. doi: 10.1016/j.socscimed.2018.10.013 [PubMed: 30412922]
- Smith J (2012). Macrostructure from Microstructure. Sociological Methodology, 42(1), 155–205. doi: 10.1177/0081175012455628 [PubMed: 25339783]
- Smith K, & Christakis N (2008). Social Networks and Health. Annual Review Of Sociology, 34(1), 405–429. doi: 10.1146/annurev.soc.34.040507.134601
- Snijders T (2011). Statistical Models for Social Networks. Annual Review Of Sociology, 37(1), 131– 153. doi: 10.1146/annurev.soc.012809.102709
- Takagi D, Yokouchi N, & Hashimoto H (2020). Smoking behavior prevalence in one's personal social network and peer's popularity: A population-based study of middle-aged adults in Japan. Social Science & Medicine, 260, 113207. doi: 10.1016/j.socscimed.2020.113207 [PubMed: 32712558]
- Valente TW, & Vlahov D (2001). Selective risk taking among needle exchange participants: implications for supplemental interventions. American Journal Of Public Health, 91(3), 406–411. doi: 10.2105/ajph.91.3.406 [PubMed: 11236405]
- Valente TW, & Saba W (1998). Mass Media and Interpersonal Influence in a Reproductive Health Communication Campaign in Bolivia. Communication Research, 25(1), 96–124. doi: 10.1177/009365098025001004
- Valente TW, Hoffman B, Ritt-Olson A, Lichtman K, & Johnson C (2003). Effects of a Social-Network Method for Group Assignment Strategies on Peer-Led Tobacco Prevention Programs in Schools. American Journal Of Public Health, 93(11), 1837–1843. doi: 10.2105/ajph.93.11.1837 [PubMed: 14600050]
- Valente TW (2010) Social networks and health: Models, methods, and applications. (1st ed.). New York, NY: Oxford University Press.
- Valente TW (2012). Network Interventions. Science, 337(6090), 49–53. doi: 10.1126/science.1217330 [PubMed: 22767921]
- Wasserman S, & Faust K (1994). Social Network Analysis: Methods and Applications Cambridge, UK: Cambridge University Press.
- Wellman B (2007). Challenges in Collecting Personal Network Data: The Nature of Personal Network Analysis. Field Methods, 19(2), 111–115. doi: 10.1177/1525822×06299133
- Wellman B, & Frank K (2017). Network Capital in a Multilevel World: Getting Support from Personal Communities. Social Capital, 233–273. doi: 10.4324/9781315129457-10

# A. Ego Network Matrix of Ties

	EGO	A1	A2	A3
EGO		0	1	2
A1	1	-	0	0
A2	1	0	-	0
A3	2	0	2	-

B. Graphic Representation of the Ego Network Matrix



# Figure 1.

Illustration of Social Network Measure Definitions and Analytic Terminology. **A.** Ego Network Matrix of Ties. **B.** Graphic Representation of the Ego Network Matrix. *The ego is labeled EGO and is demarcated with a ring around the node. Nodes beginning with A - such as A1, A2 and A3 - are alters. The arrows symbolize directed relationships, with a two-sided arrow indicating a reciprocal relationship. Links with greater width indicate greater link strength. This ego network has the following descriptive network characteristics: Size = 4, Density = 3, Out Degree = 3, In Degree = 2, Density = 4/6 = 0.67, Triad = Ego-A2-A3.* 



#### Figure 2.

Illustration of the Association between Social Influence and Child Dental Utilization Using Two Egocentric Social Influence Exposure Variables: Network Similarity and Network Density. A,B. Network visualization of network similarity between dental utilization among mothers' alters and dental utilization for the mother's child. Network similarity is defined as the proportion of alters who saw a dentist in the past year (filled blue circle) and depicted for four mothers whose child had dental utilization in the past year (A) compared to four mothers whose child did not have dental utilization in the past year (B). Each network image is for a hypothetical mother, or ego, who is denoted with a double circle. In this hypothetical example, mothers with child dental utilization had more dental utilization among their connections (greater network similarity) compared to mothers with no child dental utilization (Mean alter dental utilization 80% and 62%, respectively). C,D. Network visualization of maternal ego network density and dental utilization for the mother's child. Ego network density is the percentage of all possible relationships within each mother's social network and is illustrated for four mothers whose child had dental utilization in the past year (C) compared to four mothers whose child did not have dental utilization in the past year (D). Each network visualization depicts the links, or connections, between alter nodes for a hypothetical mother, or ego, who is not visualized. The width of the link indicates the link strength between alters. In this hypothetical example, mothers with child dental utilization had less connected networks (lower network density) compared to mothers with no child dental utilization (Mean density 0.52 and 0.88, respectively).

#### Table 1.

#### Glossary of Social Network Analysis Terminology

Definition Term Example Basics of Social Network Analysis A network member that is a distinct individual entity such as a Actor person, group, or organization Also termed Agent The focal actor of interest within an egocentric network from whose In Figures 1B, 2A and 2B, the ego is indicated Ego perspective all relational ties are defined by the double circle node E. Alter All others within an egocentric network to which the ego has a In Figure 1B, the three alters are indicated by relational tie the single circle nodes A1, A2, and A3. In Figures 2A and 2B, alters are indicated by a filled or empty circle. Relational Tie A relationship or connection between actors in a network. Relational ties may be formal (e.g., receives funding from) or informal (e.g., is friends with) and may carry characteristics of the relationship. The set of all actors and the relational ties among them Network Sociocentric A network consisting of the relational ties among all actors in a A researcher may construct a sociocentric Network single, bounded community. network of a school by asking all students to Also termed Whole Network identify other students with whom they are friends and mapping these friendship relations. Egocentric A network defined from the perspective of a single focal actor of A researcher may construct an egocentric Network interest. This includes relational ties connecting the selected actor of network by asking current smokers (egos) to interest (ego) to all known others (alters) with whom a relationship identify all of their alters who they talk to exists. The relational ties between all alters within the network are about quitting smoking. Next, the egos are also included as defined from the ego's perspective. presented with each pair of alters that they Also termed Personal Network. talk to about quitting smoking and asked if the alters talk to each other. In this way, an egocentric smoking cessation information network is constructed for each ego in the sample. Representing Social Network Data Matrix Numerical display of the relations between all pairs of actors within a Figure 1A displays the actor by actor matrix network. May be binary ("0" for no tie, "1" for tie) or valued within a used to construct the network graph in Figure range ("1" for weakest tie to "5" for strongest tie). 1b. Notice that the relations are represented with values ranging from 0 to 2, where increasing numbers indicate stronger ties. Visual model of a network including all actors (as nodes) and Graph Figure 1B displays the network graph relational ties (as links) constructed from the corresponding matrix in Figure 1A In Figure 1B, all circles (labeled E, A1, A2, Node The graphic representation of an actor within a network. Also termed Point or Vertex and A3) are nodes. Link The graphic representation of a relational tie within a network. Links In Figure 1B, all lines between nodes are links. may represent characteristics of the relational tie such as the strength Notice that increasing strength is represented by line width and direction is represented by or direction of the connection. Also termed Line, Tie, Arc, or Edge single or double-headed arrows. The link from E to A1 is unidirectional. The link between E and A2 is bidirectional. Route connecting a pair of nodes within a network graph along the Path ties between them and including any intermediary nodes Length The number of links contained within a path Distance The length of the shortest path between any pair of nodes Also termed Geodesic Distance Units of Measurement within Social Network Analysis Individual A single node Isolate A node which has no connections to any other node

Term	Definition	Example
Bridge	A link which serves to connect what would otherwise be disconnected nodes or subgroups of nodes	
Subgroup	Any subset of nodes and the links between them in which all nodes are connected directly or indirectly by at least one link Also termed <i>Component</i>	
Dyad	Two nodes and the possible links between them	In Figure 1B, nodes E-A1 form a dyad
Triad	Three nodes and the possible links between them	In Figure 1B, nodes E-A2-A3 form a triad
Entire Network	All nodes and ties within the analysis	
Unique Concepts for	Social Network Analysis in Health Research	
Access	The ability of an actor to receive or transmit an element (e.g., ideas, support, resources, health behaviors, health services, disease) as facilitated by its direct and indirect connections to other actors in the network	
Brokerage	The ability of an actor to influence or be influenced by elements as a result of their position along key paths of transmission	
Social Influence	The process by which an actor is influenced by the network due to the characteristics and structure of the actors and ties to which it is connected	

Note: Definitions adapted from Wasserman & Faust, 1994; Hawe et al., 2004; Altman et al., 2018

## Table 2.

## Network Measure Definitions

Network Measure	Definition	Example
Size	The total number of nodes in a network	In Figure 1B, the network size is 4.
Centrality	Measures of the prominence of a node within a network as determined by how extensively connected that node is to other nodes within the network	
Degree	The total number of nodes directly connected to the ego	In Figure 1B, the degree is 3.
In degree	The sum of the connections leading to an ego from their alters	In Figure 1B, the in degree is 2.
Out degree	The sum of the connections leading from an ego to their alters	In Figure 1B, the out degree is 3.
Closeness	The inverse sum of the distance in links connecting the ego to all other nodes in the network	
Betweenness	Extent to which the ego connects pairs of other nodes by falling on the shortest path between these nodes	
Structural holes	Gaps between two nodes without a direct connection and which have nonredundant other connections	
Cohesion	Interconnectedness of nodes within the network	
Density	The number of links present in a network divided by all possible links which could exist in the network	In Figure 1B, there are 4 links present out of 6 possible links. There is no direct link between A1-A2 or A1- A3. The density is 4/6 or 0.67. In Figure 2C and 2D, the number of links present is divided by the possible number of links for each of the 8 mothers. In Figure 2C, for Mom 1, this is 5/10 or 0.50.
Link strength	A value measure assigned to a link based on the reported intensity, duration, or frequency of contact between the two nodes	In Figures 1B, 2C and 2D, wider lines indicate greater link strength. In Figure 1B, the Ego-A1 and Ego-A2 links have weaker link strength and therefore thinner lines compared to the Ego-A3 and A2-A3 links.
Reciprocity	The percentage of links in a network that are bidirectional	In Figure 1B, 2 out of 4 links are bidirectional. The reciprocity is 1/2 or 0.50.
Transitivity	The probability that two nodes connected to a common node are also connected directly themselves	
Network similarity	The proportion of an ego's alters that share a characteristic of interest	In Figure 2A and 2B, the network similarity measure is the proportion of mothers' connections who saw a dentist in the past year, which is the summed number of connections who reported visiting a dentist divided by the total number of connections. In 2A, for Mom 1, this is 4/5 or 0.80.

Note: Definitions adapted from Wasserman & Faust, 1994; Hawe et al., 2004; Altman et al., 2018

## Table 3.

Example of Best Practices for Modeling Egocentric Social Network Data and Health Outcomes to Examine the Association Between Two Social Influence Variables, Network Similarity and Network Density, On Child Dental Utilization.

Best Practice Step	Example #1: Network Similarity	Example #2: Network Density		
1. Model Selection	The main outcome variable was child dental utilization in the past year (yes/no). Because the main outcome variable was binary, we used a logistic regression model.			
2. Social Network Exposure Variable(s) & Selection Considerations	<ul> <li>In this example, we chose a <b>network similarity</b> measure to describe social influence, operationalized as the proportion of mothers' alters who saw a dentist in the past year.</li> <li>We hypothesized that mothers with more dental utilization in their social networks would have a greater odds of child dental utilization compared to mothers with less dental utilization in their social networks. This research question was rooted in the phenomenon of cultural norms, such that mothers with a cultural norm of routine dental care among their alters would experience social pressure to conform and take their children to a dentist.</li> </ul>	<ul> <li>In this example, we chose a <b>network density</b> measure to describe social influence, defined as the percentage of all possible relationships within each mother's social network.</li> <li>We hypothesized that mothers with few friendships among their connections (less dense networks) would have a greater odds of taking their child to a dentist compared to mothers whose friends were all friends with each other (more dense networks).</li> <li>This research question was rooted in the phenomenon that non-redundant relationships can be beneficial because they avoid social pressure to perform undesired health behaviors within an insular environment and allow for a greater variety of ideas that influence dental health-seeking behaviors for children.</li> </ul>		
3. Covariate Selection: Sociodemographic & Health Characteristics	Sociodemographic and health covariates were selected based on previous literature about child dental utilization: child age, child dental insurance, child oral health status, mother's education and family income (Pullen et al., 2018; Fisher-Owens et al., 2007).			
4. Covariate Selection: Social Network Characteristics	<ul> <li>Mothers' network covariates were selected based on theory related to the exposure variable and child dental utilization (Pullen et al., 2018; Fisher-Owens et al., 2007):</li> <li>1. network size (count),</li> <li>2. network link strength (continuous mean relationship strength score from 1 to 10, divided by 10),</li> <li>3. network child dental knowledge (continuous proportion of alters with high dental knowledge),</li> <li>4. network dental trust (continuous proportion of alters who trust dental providers).</li> </ul>			
5. Analytic Considerations	We examined multivariable models for collinearity.			