Introducing SoNHR – Reporting guidelines for social networks in health research

Douglas A. Luke\textsuperscript{1*}, Edward Tsai\textsuperscript{2}, Bobbi J. Carothers\textsuperscript{1}, Sara Malone\textsuperscript{3}, Beth Prusaczyk\textsuperscript{4}, Todd B. Combs\textsuperscript{1}, Mia T. Vogel\textsuperscript{5}, Jennifer Watling Neal\textsuperscript{6}, Zachary P. Neal\textsuperscript{6}

Affiliations:

\textsuperscript{1}Center for Public Health Systems Science, Washington University in St. Louis, St. Louis, MO, United States of America

\textsuperscript{2}Office of Community Engagement and Health Equity, University of Illinois Cancer Center, University of Illinois-Chicago, Chicago, IL, United States of America

\textsuperscript{3}Department of Surgery, School of Medicine, Washington University in St. Louis, St. Louis, MO, United States of America

\textsuperscript{4}Institute for Informatics, Data Science, and Biostatistics, School of Medicine, Washington University in St. Louis, St. Louis, MO, United States of America

\textsuperscript{5}Brown School, Washington University in St. Louis, St. Louis, MO, United States of America

\textsuperscript{6}Department of Psychology, Michigan State University, East Lansing, MI, United States of America

* Corresponding author

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Objective

The overall goal of this work is to produce a set of recommendations (SoNHR – Social Networks in Health Research) that will improve the reporting and dissemination of social network concepts, methods, data, and analytic results within health sciences research.

Methods

This study used a modified-Delphi approach for recommendation development consistent with best practices suggested by the EQUATOR health sciences reporting guidelines network. An initial set of 28 reporting recommendations was developed by the author team. A group of 67 (of 147 surveyed) experienced network and health scientists participated in an online feedback survey. They rated the clarity and importance of the individual recommendations, and provided qualitative feedback on the coverage, usability, and dissemination opportunities of the full set of recommendations. After examining the feedback, a final set of 18 recommendations was produced.

Results

The final SoNHR reporting guidelines are comprised of 18 recommendations organized within five domains: conceptualization (how study research questions are linked to network conceptions or theories), operationalization (how network science portions of the study are defined and operationalized), data collection & management (how network data are collected and managed), analyses & results (how network results are analyzed, visualized, and reported), and ethics & equity (how network-specific human subjects, equity, and social justice concerns are reported).
present a set of exemplar published network studies which can be helpful for seeing how to apply the SoNHR recommendations in research papers. Finally, we discuss how different audiences can use these reporting guidelines.

Conclusions

These are the first set of formal reporting recommendations of network methods in the health sciences. Consistent with EQUATOR goals, these network reporting recommendations may in time improve the quality, consistency, and replicability of network science across a wide variety of important health research areas.
Introducing SoNHR – Reporting Guidelines for Social Networks in Health Sciences

Research

Introduction

Despite the dominance of the medical model of health in the 20th and 21st centuries[1], we have always understood that health is very much socially determined. Family and peer influences on smoking and diet; the role of social support on longevity and quality of life; social class and income inequality influencing access to health care; social isolation as a risk factor for depression and suicide; relational and structural factors shaping the course of pandemics – these are just some of the examples of how social factors are involved with individual, community, and population health [2].

Network science is the use of relational and structural theories to study network representations of complex social and physical systems. Although the roots of network science go back over a hundred years, its application within the health sciences is more recent. Driven by theoretical advances [3], modern computational power, and the increased availability of socially structured health data [4], the application of network designs and analytic methods within the health sciences has increased dramatically in the past few decades [5]. For example, Figure 1 shows the increase in network analytic studies over the past 20 years— with health-related network studies currently accounting for as much as 26% of all network publications.

Fig 1. While network analysis studies have grown in general over the past two decades, the percentage of those specifically from health fields has grown from just 3% in 2002 to 26% in 2021.
Modern health science has benefited greatly from the reporting guidelines movement, where methodological, analytical, and reporting best practices are used to improve research quality, validity, and impact [6,7]. Numerous reporting guidelines exist across the social and health sciences that help guide practice and reporting of clinical trials (i.e., CONSORT), systematic reviews (PRISMA), epidemiologic studies (STROBE), and implementation science studies (StaRI), to name just a few. However, until now there has not been a similar set of guidelines for reporting the results of network studies in the health sciences. This is an important gap to fill, most directly because of the prevalence of network studies, as suggested above [8]. However, a set of network reporting best practices will also benefit a notable training gap among network analysis professionals. Training in network science and analytic methods is still not regularly featured in most medical schools, schools of public health, and even many social science departments.

The overall goal of this work is to produce a set of recommendations that will improve the reporting and dissemination of social network concepts, methods, data, and analytic results within health research. These recommendations, called Social Networks in Health Research (SoNHR), focus specifically on reporting, not on how to do better network science. There already exist many excellent network analysis texts, for example, those by Scott (2017) [9] or Borgatti, Everett, and Johnson (2018) [10]. And although the primary emphasis is on reporting network studies in peer-reviewed scientific publications, we anticipate that these recommendations will be useful in other contexts including
evaluation reporting, and in training and teaching. The paper is structured in three parts: first, we describe the process we took for developing the network reporting guidelines. Second, we present and discuss the recommendations themselves. We then conclude with some recommendations on how to use the recommendations in different contexts and present some helpful resources including examples of the recommendations in practice.

Methods

Overview

This is a guidelines development study with the main goal of producing a set of reporting recommendations for studies using network data and methods in the health sciences. The general methodological approach we used for developing these recommendations was a modified expert panel consensus process, as recommended by EQUATOR (Enhancing the QUAlity and Transparency Of health Research; https://www.equator-network.org/) and implemented by other reporting guidelines development teams (e.g., Pinnock, et al., 2017) [11]. This study was approved by the Institutional Review Board of Washington University in August 2021 (ID: 202108053).

Development of initial set of recommendations and expert panel survey

An initial set of network reporting recommendations was developed by the author team, based on an informal review of the network science and health science literatures, as well as their experience using
network methods. Collectively, the author team has decades of experience conducting network studies, disseminating network studies to health and social science audiences, teaching and training students in network methods, publishing network methods texts, and reviewing grants and manuscripts using network and systems science methods. In this development phase, we produced a set of 28 candidate recommendations, grouped into five categories: 1) conceptualization, 2) operationalization, 3) data collection & management, 4) analyses & results, and 5) ethics & equity. These categories broadly represent the phases of empirical research, along with an overarching category focusing on ethical and equity aspects of network science [12].

Originally, we planned to convene network science experts to provide input and feedback on the recommendations. However, because of COVID-19 we switched to an online survey approach. The survey was developed in Qualtrics and focused on two types of feedback. Usability ratings were assessed using two items: importance of the recommendation, and clarity of the recommendation.

Importance was defined as “the degree to which you feel that the particular guideline is critical for researchers to understand and follow in their empirical studies which apply SNA methods” and was rated on a 1 (Not at all Important) to 5 (Very Important) scale. Clarity was defined as “the degree to which the wording of the guideline clearly communicates to researchers using these reporting guidelines the type of information needed to satisfy the guideline” and was rated on a 1 (Not at all Clear) to 5 (Very Clear) Likert scale.

In addition to these quantitative items, participants were asked to respond to a series of open-ended questions focusing on general improvements to the guidelines, suggestions for disseminating the guidelines, and how they might use the guidelines in their own work. Appendix A contains the expert panel feedback survey, including the initial set of 28 recommendations.
Participants

Our goal was to identify individuals who were experts in social network analysis applied to health contexts. To that end, we identified five stakeholder groups that should be represented: experts in applied health research, social network methodologists, network methods instructors, journal editors, and funders. Using these groups as guiding principles, we identified individuals who authored well-known papers, books, and analysis software; colleagues from federal funding agencies (NIH, CDC, etc.); participants in network analysis trainings (e.g., Systems Science for Social Impact); and members of the International Network for Social Network Analysis (INSNA). We invited a total of 147 people to participate, of whom 67 responded for a response rate of 45.6%. The survey ran from December 2021 through April 2022. Responses were treated as confidential; specific identifying information was stripped from the data files prior to analyses.

Participant characteristics are presented in Tables 1 and 2. The participants as a group were experienced network and health scientists, working across a wide variety of fields and research areas. Over 80% of the participants had at least six years of experience using network designs and methods, with over 20% having at least two decades of experience. Nearly all participants had published network papers and reviewed network-focused papers or grants during the past five years (94% and 88%, respectively). The participant group included ten members who have worked in an organization that funds network research activities.
Table 1. Professional characteristics of expert panel members. (Percentages can add up to more than 100% because of multiple choice options.)

<table>
<thead>
<tr>
<th>Network Research Experience</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>20+ years</td>
<td>14</td>
<td>21.5</td>
</tr>
<tr>
<td>11-20 years</td>
<td>22</td>
<td>33.8</td>
</tr>
<tr>
<td>6-10 years</td>
<td>18</td>
<td>27.7</td>
</tr>
<tr>
<td>1-5 years</td>
<td>11</td>
<td>16.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Network Science Activities (past 5 years)</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Published paper</td>
<td>61</td>
<td>93.8</td>
</tr>
<tr>
<td>SNA professional organization</td>
<td>58</td>
<td>89.2</td>
</tr>
<tr>
<td>Reviewed SNA papers/grants</td>
<td>57</td>
<td>87.7</td>
</tr>
<tr>
<td>Taught SNA</td>
<td>48</td>
<td>73.8</td>
</tr>
<tr>
<td>Worked in org. funding SNA</td>
<td>10</td>
<td>15.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Primary Disciplinary Focus</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public health</td>
<td>38</td>
<td>58.5</td>
</tr>
<tr>
<td>Sociology</td>
<td>29</td>
<td>44.6</td>
</tr>
<tr>
<td>Statistics, Math, CS</td>
<td>13</td>
<td>20.0</td>
</tr>
<tr>
<td>Organizational research</td>
<td>13</td>
<td>20.0</td>
</tr>
<tr>
<td>Psychology</td>
<td>11</td>
<td>16.9</td>
</tr>
<tr>
<td>Medicine</td>
<td>8</td>
<td>12.3</td>
</tr>
<tr>
<td>Social Work</td>
<td>7</td>
<td>10.8</td>
</tr>
<tr>
<td>All others (15)</td>
<td>31</td>
<td>47.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Health Science Applications</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implementation science</td>
<td>25</td>
<td>38.5</td>
</tr>
<tr>
<td>Health disparities, equity, justice</td>
<td>25</td>
<td>38.5</td>
</tr>
<tr>
<td>Community-focused</td>
<td>22</td>
<td>33.8</td>
</tr>
<tr>
<td>Methods development</td>
<td>21</td>
<td>32.3</td>
</tr>
<tr>
<td>Program evaluation</td>
<td>19</td>
<td>29.2</td>
</tr>
<tr>
<td>Health services</td>
<td>18</td>
<td>27.7</td>
</tr>
<tr>
<td>Mental health</td>
<td>17</td>
<td>26.2</td>
</tr>
<tr>
<td>Epidemiology</td>
<td>16</td>
<td>24.6</td>
</tr>
<tr>
<td>Health promotion</td>
<td>15</td>
<td>23.1</td>
</tr>
<tr>
<td>Addiction, substance use</td>
<td>14</td>
<td>21.5</td>
</tr>
<tr>
<td>Health policy</td>
<td>13</td>
<td>20.0</td>
</tr>
<tr>
<td>Children, youth, families</td>
<td>13</td>
<td>20.0</td>
</tr>
<tr>
<td>Older adults</td>
<td>12</td>
<td>18.5</td>
</tr>
<tr>
<td>Cancer</td>
<td>11</td>
<td>16.9</td>
</tr>
<tr>
<td>Infectious disease</td>
<td>11</td>
<td>16.9</td>
</tr>
<tr>
<td>All others (17)</td>
<td>67</td>
<td>101.5</td>
</tr>
</tbody>
</table>
Table 2. Demographic characteristics of expert panel members.

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender Identity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Woman</td>
<td>41</td>
<td>63.1</td>
</tr>
<tr>
<td>Man</td>
<td>23</td>
<td>35.4</td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>1</td>
<td>1.5</td>
</tr>
<tr>
<td>Identify as Transgender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>64</td>
<td>98.5</td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>1</td>
<td>1.5</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>53</td>
<td>81.5</td>
</tr>
<tr>
<td>Asian</td>
<td>7</td>
<td>10.8</td>
</tr>
<tr>
<td>Black or African American</td>
<td>1</td>
<td>1.5</td>
</tr>
<tr>
<td>Other race not listed</td>
<td>2</td>
<td>3.1</td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>4</td>
<td>6.2</td>
</tr>
<tr>
<td>Hispanic/Latinx Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>59</td>
<td>90.8</td>
</tr>
<tr>
<td>Yes</td>
<td>6</td>
<td>9.2</td>
</tr>
</tbody>
</table>

Development of final set of network reporting recommendations

Data from the expert panel survey were summarized and analyzed. Specifically, descriptive statistics and distributional characteristics of the importance and clarity items were examined to identify recommendations that could be dropped or recommendations that needed improvement. Thematic summaries were prepared for the open-ended questions and were also used to help differentiate between strong and weaker recommendations.
Results

Recommendations development

The quantitative results of the preliminary network reporting recommendations from the Delphi survey are presented in Table 3. In general, the preliminary recommendations received high scores on importance and clarity. The operationalization recommendations scored the highest, with average importance ratings of 4.7 (out of 5) and clarity ratings of 4.6. The five preliminary visualization recommendations scored lowest in importance (mean = 3.5). Closer examination of the visualization ratings suggested that participants fell into one of two categories—some felt these recommendations were quite important, but a substantial number clearly felt that visualizations in general were less important in network science. The detailed, item-level survey responses (including dot-plots) are available in Appendices B and C.

Table 3. Summary of importance and clarity ratings of preliminary set of network reporting recommendations.

<table>
<thead>
<tr>
<th>Domain</th>
<th>No. of Recs.</th>
<th>Importance</th>
<th>Clarity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>Range</td>
</tr>
<tr>
<td>Conceptualization</td>
<td>3</td>
<td>4.3</td>
<td>4.1 – 4.6</td>
</tr>
<tr>
<td>Operationalization</td>
<td>4</td>
<td>4.7</td>
<td>4.4 – 4.9</td>
</tr>
<tr>
<td>Data Collection &amp; Management</td>
<td>6</td>
<td>4.3</td>
<td>4.0 – 4.6</td>
</tr>
<tr>
<td>Analysis &amp; Results - Description</td>
<td>2</td>
<td>4.3</td>
<td>4.2 – 4.4</td>
</tr>
<tr>
<td>Analysis &amp; Results - Visualization</td>
<td>5</td>
<td>3.5</td>
<td>3.1 – 3.8</td>
</tr>
<tr>
<td>Analysis &amp; Results - Modeling &amp; Simulation</td>
<td>4</td>
<td>4.4</td>
<td>4.2 – 4.7</td>
</tr>
<tr>
<td>Ethics &amp; Equity</td>
<td>4</td>
<td>4.2</td>
<td>3.8 – 4.6</td>
</tr>
</tbody>
</table>
The responses to the open-ended questions indicated general enthusiasm for the network reporting recommendations (e.g., “These are awesome and well needed. Bravo!”), but more importantly included many suggestions on how to improve individual recommendations as well as their overall structure. Participants also provided numerous suggestions on how to disseminate the guidelines as well as how they might use the recommendations in their own work (see Discussion below).

Taking the importance, clarity, and qualitative feedback into account, the team revised the recommendations by dropping, rewriting, and reordering items. Items were dropped if they received lower importance or clarity ratings, or if on further reflection they seemed to be redundant with other items or were too narrow in scope. For example, one of the preliminary 28 guideline recommendations that was dropped was, “For valued networks, describe reconciliation of conflicting values when provided by both members of the dyad.” This item was overly complicated (receiving the lowest clarity score in its group) and it applies to relatively few types of network studies. The main structural change we made was to combine two separate recommendation sections (Network Description and Network Visualization) into one section.

**Final reporting guidelines for social networks in health research (SoNHR)**

Table 4 presents our final set of 18 recommendations for reporting social network methods and results, particularly within health sciences research (SoNHR). These recommendations are not a formal
checklist; we do not expect all these recommendations to apply to every specific network paper. However, the use of these recommendations will help ensure that network methods and results are reported as clearly as possible, benefiting future research.
Table 4. Final reporting guidelines - Social Networks in Health Research (SoNHR).

**Conceptualization**

C1. Clearly describe how and why networks are relevant for addressing the study's research question(s).

C2. Make the value of a network analysis apparent by explaining what kind of information a network analysis tells us that a more traditional approach would not.

**Operationalization**

O1. Define the nodes to make it clear what a node represents.

O2. Define the ties so that it is clear what each type of tie represents. Make sure to indicate if ties are directed or undirected and if ties are binary or weighted.

O3. Define the boundaries of the network so that it is clear who is included and not included in the network.

O4. State clearly the basic type of network that is being analyzed (e.g., complete network, ego networks, affiliation/2-mode/bipartite networks).

**Data Collection & Management**

D1. Describe network data collection procedures and tools (e.g., surveys and software) in enough detail to support replication. When possible, provide access to all surveys, instruments, and tools used.

D2. Describe the network data used in the study, including pre-existing data sources, how the data are stored, managed, and whether/where they are publicly available.

D3. Discuss missingness in the network data, its implications, and any attempts to impute or account for missing data (e.g., rationale for requiring one or both responses when only one member of a dyad reports a relationship).

D4. Report all data transformations (e.g., aggregation of person-level to organization-level nodes, reconciliation of conflicting link values when provided by both members of a dyad, etc.).

**Analyses & Results**

**Description & Visualization**

AD1. When discussing network statistics, be clear about the unit of analysis (e.g., node, dyad, sub-network, whole network).

AD2. Report network statistics (e.g., centrality, centralization, homophily, etc.) in terms of the real-world property of the setting being assessed (e.g., what does it mean for a particular node to have high degree centrality, for a network to have high betweenness centralization, or for nodes with similar characteristics to cluster together?)

AD3. Network visualizations should clearly illustrate study findings using design principle best practices that are appropriate for network characteristics and the goals of the visualization (e.g., using node color or shape to convey categorical properties, using node size to convey quantitative properties, limited use of labels and different shapes in large networks, varying line weights or colors in small networks).

**Modeling & Simulation**

AM1. Explain the theoretical foundations that drive the model or simulation development and testing.

AM2. If statistical network models or simulations are used, clearly specify model mechanisms and outcomes (e.g., tie formation). When possible, provide reproducible statistical programming code used in the analyses to support replication.
AM3. Present some information on how well the network model fits with the observed network data and discuss any important implications of model fit.

**Ethics & Equity**

E1. Discuss how confidentiality was explained to participants and how their confidentiality was ensured, including considerations of identifiability in network visualizations and reporting. Clarify how participants understood that information could be collected about them even if they chose not to participate, or that non-participation precluded information being collected about them.

E2. Discuss any potential biases within network structures and results that may be rooted in the network study methods (e.g., failure to capture complete networks, organizational or specific group non-participation, over-representation). Think in terms of equity and social, economic, and health justice when doing so.

---

**Conceptualization**

The first two recommendations refer to the basic conceptualization of the research project as reported in the publication. Network studies are inherently structural or relational; they concern themselves with how social objects (e.g., persons, organizations) relate to one another, or how social structures shape the flow of physical or social things such as viruses, information, money, behavior, etc. [13]. The first recommendation is to make the structural and relational conceptualizations that drive the research questions clear to the reader. This provides the basic rationale for using network methods in the first place. The second recommendation is similar but focuses on the network methods themselves. How are these network methods used in the study necessary or advantageous, given the research questions driving the study?

---

**Operationalization**

The second set of recommendations is focused on operational definitions of the various pieces of the network study, i.e., who is in the network, what the ties represent, and what kind of network is being analyzed. Definitions of the nodes are usually straightforward, but they may need additional details if
the nodes are not persons (e.g., organizations), or in the case of 2-mode networks, where both types of
nodes need to be clearly defined. For example, in an infectious disease contact tracing study, one type
of node would be people who are infected, and the other type of node would be defined as a specific
physical location where people come into contact with one another [14].

Defining the network ties used in a study often requires careful attention to detail. At a minimum, the
description of network ties should include a statement of what the tie represents (e.g., contact
frequency, trust, sexual contact) and how it was measured (e.g., self-report using a 5-point scale with
each anchor point defined). This is particularly important when more than one type of tie is collected
and analyzed in the study. Moreover, once ties have been defined, it is a good practice to refer to that
type of tie specifically (e.g., ‘friendship tie’) rather than a more generic term like ‘ties,’ or ‘connections.’
Otherwise, this can lead to the impression by the reader that a network has only one type of
relationship among network members when any network contains many different kinds of social
relationships, whether measured in the study or not.

It is always important to describe who is in the network, and, conversely, who is not in the network.
Thus, the boundaries of the network should be clear, especially for complete network studies [15]. And
this leads to the last recommendation in this section, which is a clear description of the type of
network being analyzed. First, state whether this network is complete, where most or all members of a
boundary-defined network are included (e.g., ‘all cancer healthcare providers in St. Louis County); or
ego-centric, where networks are constructed from the perspective of individual members or ‘egos’
(e.g., personal support networks of people in a study of the effects of social support). In addition, it is
often important to define the network in terms of the types of ties—are the ties directed or non-
directed, and are the ties valued or binary?
Data Collection & Management

The third set of recommendations concerns reporting of data collection and management details. Network data are quite different from the data used in the vast majority of other health and social science research. They are fundamentally relational, which implies different kinds of data collection, management, and subsequent analytic practices. Therefore, authors may need to provide more information on these data management steps than they normally would. Taken together, these data management recommendations support future replication. They also suggest network-specific data issues. For example, missing data in network studies represent more serious threats to accurate interpretation and bias [16,17], so it is important to describe how missing data are handled in any network study. Furthermore, a common data issue in survey-based network analysis is when individuals in a dyad give conflicting information about a shared tie. If this is the case, then authors should report how that conflicting information is resolved [18]. In addition, enough details on any network data transformation (e.g., turning 2-mode into 1-mode network data [14]) should be presented to support full understanding and potential future replication.

Analyses & Results

The first set of recommendations under analyses and results helps ensure that readers will fully understand the results coming out of basic descriptive network analysis studies. First, make sure to clarify what the network unit-of-analysis is for any particular set of network statistics. Readers may not fully appreciate how network results can focus on individual network members, pairs of network members, subnetworks, or complete network characteristics. Second, it is very important to convey how a particular network statistic measures or captures an underlying behavioral, structural, or...
relational characteristic of the network or network member. For example, in a communication network study, do not simply report the technical formula for a statistic such as network centralization, but describe how this measure assesses the extent to which the communication structure is more or less hierarchical in nature. Finally, many network studies employ visualizations, so it is important to design them in such a way that it is clear how the visualization supports or reveals the underlying network characteristic or analytic result. This is likely to require both general information visualization skills (e.g., Tufte, 2001), as well as network-specific graphic design principles (e.g., Shneiderman & Dunne, 2012; Ognyanova, 2021).

The second section of recommendations under analyses and results refers to good practices for reporting results from network modeling or simulation studies. These types of studies move beyond simple network descriptions and visualizations to using data and theory to pose and test hypotheses about network structures and influences. Taken together, these three recommendations will help ensure that readers will understand the theory or framework that drives the network modeling analyses, will have enough technical information to support further research replication using these models, and appreciate how well the network model or simulation fits with the observed network data [22].

Ethics & Equity

The final set of recommendations refers to special considerations around ethics and equity in network studies. Standard non-network surveys and similar methods of data collection allow for anonymity and confidentiality. However, when collecting data on networks anonymity is not possible, as the individuals or organizations in the network must be named to create ties in subsequent analyses. Even
if a person or organization chooses not to participate in a network study, they might be named by others during data collection. Offering informed consent or even truly informed consent for participants in network studies is a promising practice to ensure confidentiality [12,23]. As part of this process, those choosing not to participate should be aware that they may be named by others. Omitting non-consenters in a network study is dangerous and can greatly impact network structures, analyses, results, and interpretations [24]. More appropriate strategies for maintaining confidentiality are to anonymize data as soon as possible after data collection and to reduce the number of people who have access to the data [23,25].

Equity is an essential consideration in all scientific research, and in network analysis its importance is amplified. As stated above, maximizing response rates and capturing complete networks are also especially important in network studies. It is important to identify all who influence or are affected by the substantive area of the network under consideration, including historically underrepresented types of network members such as patients in individual networks or relatively small agencies in organizational networks. Researchers should monitor recruitment and look for patterns in non-responses. For example, one study found that network studies of sex workers commonly underrepresented men in their roles as sex workers and designed a purposeful sampling strategy more inclusive of them [26]. Similar to the ethical concerns around reporting incomplete network study findings, the underrepresentation of certain groups or organizations can distort results, or even worse, exacerbate existing inequities through exclusion.
Discussion

The SoNHR (Social Networks in Health Research) guidelines were developed to promote the clear and comprehensive reporting of network studies in the literature among health researchers. Already a well-established theoretical and methodological frame, the use of network science in health research continues to increase. However, without reporting guidelines, the utility, replicability, and impact of network science studies are diminished. These guidelines were created through an iterative, expert consensus building process and the resulting recommendations can be applied broadly across health research studies, including those in the clinical, social science, community, and population/public health research fields.

Suggested audiences and supporting resources

These guidelines have the potential to benefit at least five distinct audiences – instructors, authors, journal editors, reviewers, and readers – that capture the full process of scientific knowledge generation and dissemination.

First, these guidelines offer authors recommendations about how to report the design and results of their health-related social network research clearly and consistently. We anticipate this is especially helpful and important because social network methods are both diverse and rapidly developing, which can lead to confusion among authors about what essential aspects to report. Although the guidelines themselves offer specific recommendations, Table 5 provides examples of published research that have followed these recommendations, and thus that can serve as exemplars for authors. Although not
exhaustive, these examples represent a collection of “best practices” for each of the recommendations and illustrate how each of the recommendations can be put into action.
Table 5. Non-exhaustive list of published network studies demonstrating network reporting recommendations.

<table>
<thead>
<tr>
<th>Rec.</th>
<th>1st Author &amp; Year</th>
<th>Title</th>
<th>Description</th>
<th>Pointer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><strong>Conceptualization</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>DeLay (2017) [34]</td>
<td>Assessing the impact of homophobic name calling on early adolescent mental health: A longitudinal social network analysis of competing peer influence effects.</td>
<td>Provides a theoretical rationale for the relational nature of homophobic name calling and for studying network selection and influence processes.</td>
<td>Page 957</td>
</tr>
<tr>
<td>C2</td>
<td>Valente (2019) [36]</td>
<td>Effects of a social-network method for group assignment strategies on peer-led tobacco prevention programs in schools.</td>
<td>Links the social science literature on peer influence to the rationale for the network intervention.</td>
<td>Page 1837, last paragraph</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Operationalization</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O1</td>
<td>Lee (2011) [37]</td>
<td>Social network analysis of patient sharing among hospitals in Orange County, California.</td>
<td>Clearly describes the nodes as hospitals and providers characteristics of these hospitals.</td>
<td>Last paragraph of Background and first paragraph of Methods.</td>
</tr>
<tr>
<td>O1</td>
<td>Fuller (2007) [38]</td>
<td>Use of social network analysis to describe service links for farmers' mental health.</td>
<td>The nodes are described as “agencies known to have contact with” farming families about mental health and wellbeing in a specific town.</td>
<td>Second paragraph in Methods.</td>
</tr>
<tr>
<td>O2</td>
<td>Koehly (2003) [39]</td>
<td>A social network analysis of communication about hereditary nonpolyposis colorectal cancer genetic testing and family functioning.</td>
<td>Ties were clearly defined through measures of communication, cohesiveness, affective involvement, leadership, and conflict.</td>
<td>Measures sub-section under “Materials and Methods” section.</td>
</tr>
<tr>
<td>O2</td>
<td>Bruening (2018) [40]</td>
<td>Friendship as a social mechanism influencing body mass index (BMI) among emerging adults.</td>
<td>The description effectively communicates that ties represent friendships and are binary and directed.</td>
<td>Page 4 in “Friendships” subsection of Measurements Participants subsection within the Methods section.</td>
</tr>
</tbody>
</table>
### Data Collection & Management

<table>
<thead>
<tr>
<th>Reference</th>
<th>Title</th>
<th>Details</th>
<th>Page/Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>O3 Cauchemez (2010) [41]</td>
<td>Role of social networks in shaping disease transmission during a community outbreak of 2009 H1N1 pandemic influenza.</td>
<td>Clearly describes inclusion criteria based on an outbreak investigation in Pennsylvania.</td>
<td>1st paragraph in Results section.</td>
</tr>
<tr>
<td>O4 Broccatelli (2021) [42]</td>
<td>Social network research contribution to evaluating process in a feasibility study of a peer-led and school-based sexual health intervention.</td>
<td>Defined variables of interest as the offline friendship ties among students and student’s embeddedness in online Facebook groups.</td>
<td>Second paragraph of Analytical Method section.</td>
</tr>
<tr>
<td>O4 Dhand (2019) [43]</td>
<td>Social networks and risk of delayed hospital arrival after acute stroke.</td>
<td>Links network metrics (e.g., size and connectivity) to the underlying scientific constructs (e.g., conduits for information entry, exchange, and disruption).</td>
<td>Last paragraph of Introduction.</td>
</tr>
</tbody>
</table>

**Data Collection & Management**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Title</th>
<th>Details</th>
<th>Page/Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1 Van der Gaag (2005) [44]</td>
<td>The Resource Generator: social capital quantification with concrete items.</td>
<td>Specific resource generator items and a clear definition about what counts as “knowing” a person are provided.</td>
<td>Table 2.</td>
</tr>
<tr>
<td>D1 Lorant (2017) [45]</td>
<td>Optimal network for patients with severe mental illness: a social network analysis.</td>
<td>Survey questions and response choices are clearly explained.</td>
<td>Two paragraphs under Data Collection sub-section</td>
</tr>
<tr>
<td>D2 Kim (2020) [46]</td>
<td>COVID-19 health communication networks on Twitter: Identifying sources, disseminators, and brokers.</td>
<td>Description of a pre-existing, publicly-available dataset and its existing structures is clear. Provides a good demonstration of how to sample the most important nodes from a large dataset. Discussion of how retweets and mentions are used to measure indegree and outdegree and highlight information sources.</td>
<td>Pages 131-132 “Construction of retweet and mention networks” section. Page 133 “User level” section.</td>
</tr>
<tr>
<td>D2 Ahmed (2020) [47]</td>
<td>COVID-19 and the 5G conspiracy theory: Social network analysis of Twitter data.</td>
<td>Twitter dataset is described clearly, including what hashtags were mined.</td>
<td>Two paragraphs under Methods section</td>
</tr>
<tr>
<td>D3 Gile (2017) [48]</td>
<td>Analysis of networks with missing data with application to the National Longitudinal Study of Adolescent Health.</td>
<td>Demonstrates ERGMs of a network with missing data using four approaches: all observations (as-is), complete-case (removing nodes of non-responders), incomplete-case (uses all nodes but not the links between any dyads including non-responders), and differential popularity (models differences between respondents and non-respondents to partially compensate).</td>
<td>Pages 507-517, &quot;Analyzing adolescent friendship networks&quot; section.</td>
</tr>
<tr>
<td>D3 De Moor (2020) [49]</td>
<td>Assessing the missing data problem in criminal network analysis using forensic DNA data.</td>
<td>Information on how previous research has handled missing data and how the current method handles missing data.</td>
<td>Section 4.2 under the Methods section</td>
</tr>
<tr>
<td>D4</td>
<td>Carothers (2022) [50]</td>
<td>Mapping the lay of the land: Using interactive network analytic tools for collaboration in rural cancer prevention and control.</td>
<td>Clearly indicated the unit of analysis in ERGMS: &quot;The unit of the analysis is the ordered pair of students in a classroom (xij), and the dependent variable is the observed value of a friendship or dislike tie (1 = present, 0 = absent).&quot;</td>
</tr>
<tr>
<td>D4</td>
<td>Scott (2005) [51]</td>
<td>Social network analysis as an analytic tool for interaction patterns in primary care practices.</td>
<td>Discusses aggregation of individual responses to agency-level nodes, symmetrization of non-directed networks, and reconciliation of one valued network and one directed network. Describes how direct observations were transcribed into directed and undirected adjacency matrices that then formed the basis of the network data.</td>
</tr>
</tbody>
</table>

**Analyses & Results (Description & Visualization)**

| AD1 | de la Haye (2017) [52] | The dual role of friendship and antipathy relations in the marginalization of overweight children in their peer networks: The TRAILS Study. | Clearly indicated the unit of analysis in ERGMS: "The unit of the analysis is the ordered pair of students in a classroom (xij), and the dependent variable is the observed value of a friendship or dislike tie (1 = present, 0 = absent)." | Page 3 |
| AD1 | Hawe (2008) [53] | Use of social network analysis to map the social relationships of staff and teachers at school. | Analyses are clearly described in terms of node vs. network unit if analysis, such as centralization. | Paragraphs 5-10 under Methods section. p. 76-79 |
| AD2 | Weeden (2021) [54] | Still a small world? University course enrollment networks before and during the COVID-19 pandemic. | Interpreted changes in node degree, proportion of students in largest bi-component, average distance, and within-field clustering within a university course enrollment network before and during the COVID-19 pandemic. | |
| AD3 | Block (2020) [56] | Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. | Good use of figure design (especially node color and tie bolding) to clearly communicate shortest pathways and infection risk. | Figures 1 & 2 |
| AD3 | Ahn (2011) [57] | Flavor network and the principles of food pairing. | Good use of color, node size, and tie size to designate food groups and network structures. | Figure 2 |

**Analyses & Results (Modeling & Simulation)**

| AM1 | Bond (2014) [58] | Friends or foes? Relational dissonance and adolescent psychological wellbeing. | Provides a description of ERGMs & the foundations underlying ERGM models. Figure 1 | Page 4 & Figure 1 |
## AM1 Kiuru (2012) [59]


Provides a description of what each parameter included in the model tests.

Tests two different mechanisms of socialization (contagion & convergence) and describes how different parameters (average alter & average similarity) were used in SAOMs to test each.

Page 252

## AM2 Combs (2022) [60]

Simulating the role of knowledge brokers in policy making in state agencies: An agent-based model.

“A week in the life of an agent” illustrates how opinions about health policies diffuse across the social networks in the agent-based model.

Sections 2.3.4

## AM2 Mercken (2009) [61]

Dynamics of adolescent friendship networks and smoking behavior: Social network analyses in six European countries.

Clearly lays out how the network model is able to distinguish between social influence and social selection mechanisms for adolescent smoking behavior.

Figure 1 & Table 2


Systems analysis of collaboration in 5 national tobacco control networks.

Presents model fit statistics and graphics for the main ERGM model.

Figure 2 and text on page 1294

Results section, particularly the 2nd column on page 1027

## AM3 Morris (2009) [63]

Concurrent partnerships and HIV prevalence disparities by race: linking science and public health practice.

Conducts a simulation to examine how assortative mixing drives racial disparities in HIV prevalence; compares results of the simulations to observed survey data, particularly AddHealth.

## Ethics & Equity

### E1 D'Angelo (2021) [12]

The presentation of the networked self: Ethics and epistemology in social network analysis.

‘Encounters in the field’ section contains three case studies of network analysis that help to illustrate the importance of confidentiality.

Pages 23-26

### E1 Banbury (2017) [64]

Can videoconferencing affect older people's engagement and perception of their social support in long-term conditions management: a social network analysis from the Telehealth Literacy Project.

‘Procedure’ section discusses how researchers explained confidentiality to participants and obtained informed consent.

Page 940

### E2 Bungay (2016) [26]

Addressing underrepresentation in sex work research: Reflections on designing a purposeful sampling strategy.

‘Designing a Purposeful Sampling Strategy’ section explains how the study increased inclusion of historically underrepresented network members.

Pages 969-971

### E2 Lobb (2014) [65]

Using organizational network analysis to plan cancer screening programs for vulnerable populations.

The ‘Discussion’ section explains that although community health centers are important in cancer care networks, they are often peripheral to or isolated from disconnection to provincial providers in Canada.

Pages 361-363
Second, these guidelines can be used by instructors of social network coursework to inform the topics covered in such courses, and by their students to rapidly develop an understanding of the most important features of social network research. For example, modules in an introductory course on social network analysis could roughly follow the major headings of the guidelines, while individual recommendations under each of these headings could serve as learning objectives within these modules.

Third, these guidelines can serve as a helpful checklist for reviewers when evaluating health-related social network research. Reviewers can use the guidelines and individual recommendations to ensure the completeness of their evaluations of social network studies. In addition, following these guidelines to evaluate social network research may improve the consistency between reviewers, which is often quite low [27], as they provide a common set of reporting standards to evaluate.

Fourth, these guidelines can be required as reporting standards by journal editors who publish health-related social network research to ensure the standardized reporting of social network methods and results. This type of editorial requirement is not without precedent. For example, many journal editors already require reporting guidelines developed for other methods including PRISMA for systematic reviews and meta-analyses [28], JARS-QUAL for qualitative research [29], JARS-QUANT for quantitative research [30], and CONSORT for randomized trials [31].

Finally, these guidelines have benefits for readers of health-related social network research. When implemented, these guidelines should improve the detail, clarity, and transparency of social network methods and results for readers. Additionally, when put into practice, these guidelines should ensure
that critical methodological and analytic details are provided in social network papers, making it easier to re-use published social network research in meta-analyses and systematic reviews.

**Strengths, gaps, and next steps**

This is among the first set of reporting guidelines for social network science concepts, methods, and results. Although a number of network data formatting and reporting guidelines exist (e.g., Bagrow, 2022) [32], to our knowledge there are no guidelines focusing on best reporting practices in scientific dissemination, especially within health research. These guidelines are being published in EQUATOR, which will increase their visibility and longevity and hopefully enhance their effectiveness. Finally, these reporting guidelines include equity and ethical considerations, which are important, require special consideration in network science, and have previously been underdiscussed.

Although we developed these reporting guidelines to be broadly applicable to many kinds of network studies, they may be less helpful (or need to be adapted) for some kinds of network study designs or analytic frameworks (e.g., purely qualitative network studies, ego-centric designs, dynamic analyses, etc.). Similarly, the scope of these recommendations is somewhat high-level, we deliberately left out extremely specific reporting suggestions (e.g., guidance around the use of weighted ties in network visualizations) that would not be applicable to many network studies. Finally, for the expert panel consensus process, we may have missed some important viewpoints and distinct experiences that could have helped inform the network reporting guidelines. However, the relatively large, experienced, and diverse set of participants we ended up with (see Tables 1 & 2) somewhat mitigates that concern.
We will continue to work on disseminating the SoNHR guidelines beyond EQUATOR and then hosting these guidelines on relevant websites (including at cphss.wustl.edu). We also encourage interested partners to further spread the guidelines and share how they have used them (see above). In particular, case studies of the application of the network reporting recommendations would be useful for the field of network health research.

As we gain a better understanding of how social connections, structures, and dynamics are implicated in human disease processes, effective healthcare delivery, and promotion of public health, there will be a concomitant need for appropriate scientific frameworks, study designs, and analytic approaches. Network science and social network analysis are particularly well-suited for these types of health research. The network reporting guidelines presented here will be helpful for ensuring that the knowledge generated in these studies can have the broadest impact, both scientifically and socially.
Acknowledgements

We are grateful for the contributions and expertise of our recommendations development expert panel. They were crucial for helping us develop and finalize the SoNHR recommendations.
References


9. Scott J. Social Network Analysis. 1 Oliver’s Yard, 55 City Road London EC1Y 1SP : SAGE Publications Ltd; 2017. doi:10.4135/9781529716597


research in psychology: The APA Publications and Communications Board task force report.


64. Banbury A, Chamberlain D, Nancarrow S, Dart J, Gray L, Parkinson L. Can videoconferencing affect older people’s engagement and perception of their social support in long-term conditions

Supporting information

S1 File. Appendix A.

S2 Table. Appendix B.

S3 Fig. Appendix C.

S4 File. Feedback Data

S5 File. Feedback Data Variable Dictionary