Health surveillance through social networks

Weihua An\textsuperscript{a,}, Long Doan\textsuperscript{b}

\textsuperscript{a} Indiana University, Department of Sociology and Department of Statistics, 752 Ballantine Hall, 1020 E. Kirkwood Ave., Bloomington, IN 47405, USA
\textsuperscript{b} Indiana University, Department of Sociology, 744 Ballantine Hall, 1020 E. Kirkwood Ave., Bloomington, IN 47405, USA

\section*{Article Info}

\textbf{Keywords:}
Health surveillance
Social networks
Informants
Reporting bias
Peer reports
Smoking

\section*{Abstract}

We propose a network-based method to monitor health behaviors and point out the general conditions for it to work effectively. The method helps to identify effective informants for monitoring future health behaviors and to triangulate self-reports of sensitive health behaviors. We demonstrate the method by studying the smoking behaviors of over 4000 middle school students in China. Using students' observations of their schoolmates smoking in the past 30 days, we construct smoking detection networks and examine the patterns of smoking detection. We find that smokers, optimistic students, and popular students make better informants than their counterparts. We also find that using three to four (or the 3rd quartile of) positive peer reports can uncover a good number of under-reported smokers while not producing excessive false positives.

\section*{1. Introduction}

In this paper we propose and demonstrate a network-based method of monitoring health behaviors. Specifically, we argue that collecting peer reports of health behaviors in addition to self-reports offers two major benefits for health surveillance. First, it helps to identify the characteristics of good informants who can be used for monitoring future health behaviors. Second, peer reports can help to triangulate self-reports and correct possible self-reporting bias.

To demonstrate the method, we ask over 4000 students from six middle schools in China to report the students in their schools whom they have observed smoking in the past 30 days. Using these reports, we construct smoking detection networks and examine what student characteristics are most correlated with smoking detection. We find that smokers, optimistic students, and popular students make better informants than their counterparts. We also find that using three to four positive peer reports can uncover a good number of under-reported smokers while not producing excessive false positives.

In the following sections, we outline the method of health surveillance through social networks. In our empirical analyses, we first examine the features of our network data, including friendship networks, cigarette exchange networks, and smoking detection networks. We also outline the characteristics of the best informants by comparing students' capability in detecting peers smoking. Then we use exponential random graph models (ERGMs) to examine the patterns of smoking detection and more formally identify the characteristics of the students who are more likely to detect others' smoking (as well as who are more likely to be detected). After that, we demonstrate how to use peer reports to triangulate self-reports of smoking behavior. We end by discussing the potential of this method for health surveillance research more broadly.

\section*{2. Health surveillance through social networks}

Despite the voluminous literature in social networks and health (e.g., Bearman et al., 2004; Christakis and Fowler, 2007; Cotterell, 2007; Hoffman et al., 2007; Ali and Dwyer, 2009; Cornell, 2009; Cornell and Laumann, 2011; Mercken et al., 2010; Liu et al., 2010), few studies have made the connection between networks and health surveillance. In this paper we highlight the importance of collecting peer reports of health behaviors in addition to self-reports for health surveillance. Specifically, we argue that peer reports can be used to construct health surveillance networks that can further be used to identify key informants and peer reports can also be used to triangulate self-reports and address possible self-reporting bias.

The method of key informants is widely used in criminology (Pauwels and Hardyns, 2009), but less often so in health research (but see Pai et al., 1998; Campbell et al., 2008). We argue that identifying key informants is important, because informants can be used later for multiple purposes, such as identifying subjects who regularly display a certain health behavior or monitoring the trend of the health behavior.

\* Corresponding author. Tel.: +1 812 856 1370.
E-mail addresses: weihu@indiana.edu (W. An), longdoan@indiana.edu (L. Doan).

http://dx.doi.org/10.1016/j.socnet.2015.02.001
0378-8733/© 2015 Elsevier B.V. All rights reserved.
Specifically in our case, we asked each student to report whom they have seen smoking in the past 30 days and use that information to construct smoking detection networks. By examining these networks, we can identify the characteristics of the best informants (i.e., those who are able to detect the most smokers). More embeddedness in groups of smokers may make a person a better informant, but we do not assume that highly embedded students are necessarily the best informants. Our study differs from previous ones that ask respondents to report the overall prevalence of smoking in a school (e.g., Unger and Ann Rohrbach, 2002), because we ask respondents to identify each individual smoker. It also differs from those using “student leaders” to report on the smoking prevalence (e.g., Prokhorov et al., 1993), because we do not assume that student leaders are naturally the best informants. Rather, we argue that the best informants will vary by the health behaviors and social contexts and their characteristics may be obtained by examining the health behavior detection networks. Doing so may lead to unconventional and often new sources of information. Indeed, in our case while we find that smokers make good informants for smoking behaviors, we also find that students with less intuitive characteristics (e.g., popularity and optimistic personality) also make good informants.

The health surveillance networks may also provide a cost-effective alternative or supplement to biological tests and other methods of measuring health behaviors. Conventional health surveys often suffer from self-reporting bias. For example, self-reports have led to inconsistent smoking prevalence rates within a population (Henriksen and Christine, 1999) or under-reporting of sensitive health behaviors (Kenkel et al., 2003). To increase the validity of self-reports, some have turned to longitudinal methods to check for errors and inconsistencies (Henriksen and Christine, 1999; Johnson and Mott, 2001; Mair et al., 2006), while others have turned to biological testing (Vartiainen et al., 2002). Regardless of the effectiveness of these methods, they all tend to bear a large cost. In contrast, the cost, in terms of resources as well as expertise and invasiveness, of implementing a health surveillance question is relatively small while the information provided can be used to triangulate self-reports.1 This is especially true when the health surveillance question is administered as part of an existing survey.

One empirical problem with this method is how many positive peer reports to use to verify a self-report. On one hand, using one positive peer report may lead to many false positives, as it is possible that some peers unintentionally (e.g., students mistakenly fill the surveys) or intentionally misreport others’ health behavior. Using a much larger number of positive peer reports, on the other hand, may lead to missing some subjects with a certain health behavior (i.e., retaining too many false negatives). In our study, we show that using three or four peer positive reports can help to uncover a good number of under-reported smokers while not producing excessive false positives. In practice, however, the ideal number of positive reports required to verify a self-report may vary by context and depends on the researcher’s goals.

Like other methods, the health surveillance method works better under certain conditions. Given that the method relies on peer reporting the health behavior being monitored should be publicly observable to peers. Health surveillance is also better when there is motivation to misreport one’s own behavior but there is less motivation for others to misreport the behavior. Third, health surveillance works better the smaller the cost of reporting others. This cost may include fear of retaliation or loss of confidentiality. Therefore, when the involved behavior is deviant or illegal, people may be more cautious about reporting others.

We argue that the health surveillance method is well suited for monitoring adolescent smoking. First, research has shown that smoking among adolescents tends to occur in groups or in public settings (Stewart-Knox et al., 2005; Urberg et al., 1997). Indeed, 73 percent of the self-identified smokers in our study report that they usually smoked with other students and 62 percent of them indicate that they obtain cigarettes from other students. Second, past research has shown that younger populations tend to under-report their smoking (Kenkel et al., 2003). In contrast, the motivation for them to misreport others’ smoking behaviors may be smaller, especially in cases like ours where their reports are confidential. Hence, on the one hand, we expect a significant portion of the students will not report others’ smoking because there is not much incentive for them to do so. On the other hand, we also expect a sizable portion of them will provide reports of others’ smoking. In the following, we present an empirical case of how to monitor adolescent smoking through peer reporting.

3. Data and methods

3.1. Data

Between 2010 and 2011, we conducted two waves of surveys about smoking and social networks with 4094 students from six middle schools in China. A major reason that we chose China as the field site is because of its high prevalence of smoking. Recent research shows that 66 percent of the males and 3 percent of the females above the age of 15 in China are smokers (Hu et al., 2008) and three out of five smokers start smoking as teenagers (Cheng, 1999). The schools we surveyed come from a site in central China. Although not randomly selected, the demographic and economic conditions of these schools are similar to a significant portion of Chinese middle schools.2 All data except for the cigarette exchange networks and personal factors come from the second wave, which is collected four months after the first wave.

3.2. Measures

3.2.1. Smoking status

We asked students to report whether they had smoked within the past 30 days and used a binary variable to indicate their smoking status (1 = yes; 0 = no).

3.2.2. Smoking detection networks

Students are asked to list up to four students whom they had seen smoking cigarettes within the past 30 days. Using students’ reports of other students’ smoking, we constructed a smoking detecting network for each of the six schools. Each node in the network represents a student and each link a smoking detection relationship.

3.2.3. Friendship networks

Students were asked to name up to ten of their closest friends in the school. Using these nominations, we constructed a friendship network for each school.

3.2.4. Cigarette exchange networks

At the first wave of the survey, students were also asked with which of their friends they have ever exchanged cigarettes. Using this information, we constructed a cigarette exchange network for each school.

---

1 In the discussion, we provide more detailed comparisons of our method with the biological tests.

2 Due to confidentiality, we cannot release the specific information about this research site.
3.2.5. Personal factors

At the first wave, we also asked students to report information on a list of personal factors, from which we created a binary variable for male students (1 = yes; 0 = no), an ordinal variable for academic ranking (1 = bottom 10, 2 = bottom 11–20, 3 = between top 20 and bottom 20, 4 = top 11–20, 5 = top 10), another ordinal variable for personality on a scale ranging from “very pessimistic” to “very optimistic” (1–5), and indicators on the grade and class the students belong to. We also asked students to report information about their family background, such as their father’s and mother’s education (1 = no formal education, 2 = preliminary school, 3 = middle school, 4 = high school, and 5 = junior college or above) and family’s economic status (1 = very difficult to 5 = very good).

3.3. Empirical expectations

We examine the relationships between smoking detection networks and other measures at two levels. First, we examine how personal factors affect students’ probabilities of both detecting others who are smokers and being detected as smokers. Second, we examine the effects of friendship networks and cigarette exchange networks on smoking detection, namely, whether having ties in a friendship or cigarette exchange network corresponds to having a tie in the smoking detection network.

Because smoking is much more prevalent among boys in China than girls (Yang et al., 2004), we expect boys to be more likely to be detected as a smoker. Likewise, they should also be better informants. Self-identified smokers should be more likely than non-smokers to be detected. Given past work on homophilous friend selection, we would also expect smokers to be better informants than non-smokers. Social delimitations like gender, grade, class, family background serve to constrain social relationships, and as such, we expect smoking detection to be homophilous based on these factors. Finally, other network ties are of particular interest. We expect that friends are more likely to detect a smoker than non-friends. Having exchanged cigarettes with others will also increase the likelihood of detecting their smoking additional to the effects of being friends.

3.4. Analytic strategy

Our analysis is separated into three main sections. First, we provide descriptive statistics and plots of the smoking detection, friendship, and cigarette exchange networks. This allows us to examine the features of the networks generally. We also contrast the characteristics of the best and worst informants in descriptive terms. Second, we more formally model the smoking detection networks and examine their features. Third, we show the feasibility of using health surveillance to address self-reporting bias.

To formally model the smoking detection networks, we use the Exponential Random Graph Models. In an ERGM (Wasserman and Pattison, 1996; Goodreau et al., 2008; Robins et al., 2007a,b), the probability of observing the current network, w, is assumed to be:

\[ P_r(W = w | X) = \frac{\exp(\beta g(w, X))}{K} \]

where \( W \) is a stochastic network, \( w \) represents the currently observed network, \( X \) the covariates, \( g(w, X) \) a list of network or covariate effects of interest, \( \beta \) the corresponding coefficients for these effects, and \( K \) the normalizing factor ensuring the probabilities sum to 1. Alternatively, this model corresponds to an extended logit model (Hunter et al., 2008) for observing the presence of a tie between two actors, namely,

\[ \text{logit}(w_{ij} = 1 | w^*, X) = \beta^T \delta^I(w, X) \]

where \( \delta^I(w, X) \) is the change in the covariates values and network structures (e.g., reciprocity and transitivity in forming ties) when the network is changed due to the presence of a tie \( w_{ij} \).

In the first ERGM, we examine the effects of personal factors on smoking detection. We include gender because smoking and social interactions (and so maybe detection) tend to be stratified by gender in China. We include self-reported smoking status as a way to see whether self-identified smokers are indeed more likely to be detected for smoking by others. This allows us to gauge whether health surveillance is even minimally detecting the correct students. We include personality measures as a proximate way to control for people’s tendency to report about others and to be socially engaged with others. We also include a series of measures of homophily on gender, family background, grade, class, and academic performance because social interactions (and perhaps also smoking detection) tend to be delimited by these factors. In the second ERGM, we additionally include network structure terms to account for the high number of isolates in the detection networks and to model reciprocity in smoking detection.

The relative sparseness of the detection networks does not allow for good-fitting other higher order structural terms. However, accurately modeling the structural terms may shed light on the processes through which the best informants identify smokers. In that vein, we adopt current methodological recommendations for modeling large, sparse networks and specified a third ERGM with higher order structural terms on only the non-isolates in the detection networks (Goodreau et al., 2008). Following Hunter (2007) and Papachristos et al. (2013), we model transitivity using two terms: geometrically weighted edgewise shared partners (gwesp) and geometrically dyadwise shared partners (gwdsp). The former term captures whether two nodes who have a shared partner are more likely to be connected and so to close a triangle. The latter term captures the likelihood of any two nodes, even if they do not share a tie, are connected indirectly by another node. If smoking is a group behavior and smokers truthfully report on others’ smoking, detection networks will have many closed triangles and we expect to see a positive coefficient on the gwesp term. If smoking detection is hierarchical (e.g., A reports B and B reports C, but neither A or C reports each other), we expect to see a positive coefficient on the gwdsp term. To model the skewed distribution of ties in the detection networks, we include two terms: geometrically weighted indegree (gwidegree) and geometrically weighted outdegree (gowdegree). A positive coefficient indicates tie heterogeneity, where some students are sending and receiving disproportionately more ties (i.e., detections) than others. A negative coefficient indicates tie homogeneity, or a similar tendency for students to detect others and to be detected.

Finally, we return to the main model (i.e., model 2) to add the social interaction factors from the friendship and cigarette networks. This helps to account for the multiplicity of social ties and the effects of past interactions.3

3 Past research on student reporting of peers’ behaviors indicate that social norms against “snitching” or “tattling” often result in underreporting (Knowledge Networks, 2002). But, research also suggests that guaranteed confidentiality and feelings of self- and group-efficacy increase reporting of deviant behaviors (Wilson-Simmons et al., 2006). This suggests that while we can guarantee students’ confidentiality in our study, there may be residual personality effects for which we attempt to control with these measures.

4 The way we modeled the higher order structural terms, although imperfect, allows us to capture the structural features of the connected and perhaps more interesting part of the smoking detection networks. As a result, however, the inferences we draw from doing so are only applicable to the connected parts of the networks. Given our interest in describing the detection patterns rather than the whole network, this limitation is not as problematic in this context.

5 We also fitted an ERGM with the higher order structural terms added to this model. The model was fitted on non-isolates, because as mentioned before, the
All ERGMs are fitted through the “Statnet” package in R (Handcock et al., 2003). We fit the ERGMs for each of the six schools separately and combine the results through a meta-analysis in which the original ERGM estimates in each school are weighted according to the inverse of their variances (Snijders and Baerveldt, 2003). The meta-analysis approach has been a popular method to combine results from multiple network models and has been incorporated in network analysis software like “RSiena” (Ripley et al. 2014). It allows us to present the overarching patterns of the networks succinctly. The coefficients reflect the average effects of the parameters in the networks. The associated p-values evaluate whether the average effects are statistically significant at the 5% level. We also provide testing results regarding whether the estimated coefficients are equal across all schools. In addition, we also provide the estimates for each school in the Supplementary Appendix (Table S1).

After examining the patterns of the smoking detection networks, we show the feasibility of using peer detections to triangulate self-reports. We conceptualize the triangulation in probabilistic terms. That is, we argue that the more times a self-reported nonsmoker is detected by peers for smoking, the more likely that individual has smoked. We use both sensitivity analysis and empirical analysis to show that using three to four (or the third quartile of) positive peer reports can help to uncover a good number of under-reported smokers while unlikely producing excessive false positives.

4. Results

4.1. Descriptive statistics

We begin with various descriptions of the friendship, cigarette exchange, and smoking detection networks in two example schools: school 3 and school 4. The overall patterns in other schools look similar to one of these two schools. Fig. 1 shows the sociograms of non-isolates in those networks with self-identified smokers colored in red. We can see that most of the cigarette exchanges were among smokers. But a number of students with whom others report having exchanged cigarettes claim to have never smoked. Similarly, a large number of students report being non-smokers despite having been seen smoking by others. These discrepancies indicate that some students may have under-reported their smoking status. The positioning of the nodes is fixed across networks so that it is possible to directly compare the structure of the networks. The cigarette exchange and smoking detection networks are much sparser and have smaller proportions of mutual ties than the friendship networks. In addition, the ties in the cigarette exchange and smoking detection networks seem to be a subset of those in the friendship network in each school, which suggests that both cigarette exchange and smoking detection are more likely to occur between friends. Although the overall patterns of smoking detection are similar across schools, we can also see noticeable variation in the network structures across these two schools in Fig. 1. This possibly reflects school differences in friendship and smoking norms. These differences warrant examining the network features of each school separately, as we do in the analyses.

Because the smoking detection network is our primary network of interest, we describe this network in further detail. The distributions of indegree and outdegree for non-isolates in each school are shown in online supplemental Fig. S1. A node’s indegree in the smoking detection network is the number of people who have detected that node as a smoker. A node’s outdegree is the number of smokers that a node has detected. The distribution of indegree and outdegree for non-isolates follow the same general patterns across schools. The distribution of indegree for these nodes are right skewed, with most having an indegree of one, meaning that only one other person saw that node smoking in the past 30 days. There are also students who are detected from nine times to twenty-two times. In terms of outdegree centrality, non-isolates likely name either just one or all four possible names, but a sizeable proportion of nodes also have an outdegree of two or three, which largely reduces concerns that students may not be willing to report on others.

We also compare the characteristics of the informants by how many other students they detect as smokers in Table 1. The most striking pattern we find is that the best informants (i.e., those who are able to detect the most smokers) tend to be popular smokers. Among those who detect no smokers, only 4 percent are self-identified smokers. On average, 6.68 other students nominate them as friends and they nominate 6.49 others as friends. In contrast, among those who nominate the maximum four smokers, 15 percent are self-identified smokers. On average, 7.81 other students nominate them as friends and they nominate 9.26 others as friends.

4.2. Patterns of smoking detection

Another way to examine the characteristics of the best informants is to study the patterns of smoking detection via ERGM. Moreover, the ERGM can also help us examine the characteristics of the students that are most likely to be detected as smokers. Table 2 shows the combined results of the ERGMs of smoking detection networks in the six schools. In addition to traditional tests of the significance of each variable in predicting a tie in the smoking detection network, we also include results for testing whether the estimated coefficients are different across the six schools (shown in the Q columns).

4.3. Personal factors

Model 1 identifies key node attributes that correlate with smoking detection. The odds of self-reported smokers being detected for smoking is over three times the odds of self-reported non-smokers being detected for smoking (odds ratio is e^{1.21} = 3.35, p < 0.001). Boys have about ten times higher odds of being detected for smoking than girls (odds ratio is e^{2.39} = 10.91, p < 0.001). This may indicate that boys tend to under-report their smoking status. Personality traits are also significantly related to being detected as a smoker. In general, more optimistic students are significantly less likely to be detected as a smoker than less optimistic students.

In terms of sending ties, although boys are more likely to be detected, they are not significantly more likely to detect smokers than girls. Self-identified smokers are significantly more likely to detect other smokers than non-smokers. Students with very high levels of optimism are also more likely to detect smokers than students with very high levels of pessimism. This is consistent with our argument that higher optimism reflects more social engagement and efficacy, suggesting that these students are more likely to report others’ smoking despite social pressures otherwise.

Model 1 also includes terms to capture homophily among certain types of students. Family background variables such as family’s economic status are both statistically significant and positive. This suggests that students from similar economic backgrounds are more likely to detect each other. There are also strong homophilous effects of belonging to the same grade and class. This may reflect the social constraints of spending most of their school day with students in the same grade and class. Surprisingly, we

relative sparseness of the complete networks do not allow for good-fitting the higher order structural terms. We also removed corresponding isolates in the friendship and cigarette exchange networks. The model did not converge after various attempts and re-parameterizations.
find no significant homophily by gender, suggesting that although smoking itself tends to be highly gendered in China, smoking detection is not stratified by gender.

4.4. Network structures

We add two structural terms in Model 2: mutual and isolates. We find no significant evidence that detection tends to occur mutually, but we find a significant number of isolates in this network. This may be because the majority of middle school students are not smokers. Also importantly, accounting for these structural terms makes several personal factors significant in Model 2, most notably gender. Whereas boys are more likely to be detected as smokers, they actually make worse informants. Boys have 30 percent lower odds ($e^{-0.40} = 0.67$) of detecting a smoker than girls. This may indicate that boys are less willing to report on their peers. Otherwise, the effects of the personal factors remain largely similar.

Model 3 includes additional higher order structural terms. As previously stated, the model is fitted only on the non-isolates because of the sparseness of the detection networks. The inclusion of the higher order structural terms do not substantially alter the estimates on personal factors and homophily. Some of the personality variables lose significance in the model as does homophily based on economic status. The effect sizes remain approximately similar, however, so the loss of significance here may simply be a reflection of the much reduced network size in this model. More importantly,

![Fig. 1. Example sociograms of non-isolates in friendship networks, cigarette exchange networks, and smoking detection networks for two of six schools.](image)

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Mean characteristics of student informants by students' number of detection.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of detections</td>
<td>Indegree centrality$^a$</td>
</tr>
<tr>
<td>0</td>
<td>6.68</td>
</tr>
<tr>
<td>1</td>
<td>7.26</td>
</tr>
<tr>
<td>2</td>
<td>7.49</td>
</tr>
<tr>
<td>3</td>
<td>7.09</td>
</tr>
<tr>
<td>4</td>
<td>7.81</td>
</tr>
<tr>
<td>Total</td>
<td>6.85</td>
</tr>
</tbody>
</table>

$^a$ Centrality measures are from the friendship network, representing the popularity of the student in terms of number of friends.

$^b$ Binary variables (1 = boy and smoker respectively). The summary statistics is the proportion of 1s.

$^c$ Personality is on a scale from 1 (Very pessimistic) to 5 (Very optimistic).

$^d$ Family's economic status is on a scale from 1 (Very difficult) to 5 (Very good).

$^e$ Academic ranking is on a scale from 1 (Bottom 10) to 5 (Top 10).
we find significant effects of gwesp, gwdp, and gwdegree. The negative and significant coefficient for gwdp indicates that these networks have fewer unconnected dyads with shared partners than we would expect by chance. The positive and significant gwsp term suggests that there are many closed triangles in the networks. Together these two results suggest that smoking detection tends to occur within small groups, possibly reflecting that smoking itself tends to occur in group settings. The nonsignificant coefficient for gwdegree suggests there is no clear tendency for some students to be disproportionately detected as smokers, once covariate effects and other network effects are accounted for. The negative and significant coefficient for gwdegree indicates that there is not much variation in students’ power of detecting smokers, which possibly reflects the fact that we have capped the maximal number of detection at four and so greatly limited the variation by design.

4.5. Social interactional factors

Model 4 includes friendship and cigarette exchange ties. Including these ties does little to change the effects of the personal factors and the strong presence of isolates, but these social interactional factors are themselves highly significant. The odds of detecting friends’ smoking is about five times the odds of detecting non-friends’ smoking (odds ratio is \( e^{1.58} = 4.85, p < 0.001 \)). This suggests that those students who are socially active and engaged with others are better informants about others’ smoking behavior. In contrast, in supplemental analyses, we find that students nominated by others as friends are not significantly more likely than other students to detect the nominator’s smoking (see Table S2). Together the two results indicate that detection is largely asymmetric. If a student nominates another student as a friend, the nominator is more likely to detect the nominee’s smoking, but the nominee is not more likely to detect the nominator’s smoking. This may reflect status and attention differential between the nominators and the nominees, namely, that the nominators pay more attention to what their nominees are doing than vice-versa.

Additionally, having a tie in the cigarette exchange network increases the odds of detecting a smoker by a factor of 4.18 (\( p < 0.001 \)). Both the effects of friendship and cigarette exchange suggest that smoking detection is strongly related to social interactions. In other words, social ties of one form are predictive of ties of another form.

4.6. Model diagnostics

Fit diagnostics of our example schools (school 3 and school 4) for models 3 and 4 are provided in the Supplementary Appendix (see Figs. S2 and S3). In general, the models adequately reproduce the observed networks with the exception of under-predicting the number of nodes that nominate the maximum 4 smokers and over-predicting those who nominate 3 smokers. This pattern likely indicates a priming effect of having the detection nominations capped at 4. The models were also checked for convergence based on Hotelling’s \( T^2 \) tests of equality of the MCMC simulated networks.
Table 3 Distribution of smokers and non-smokers by health surveillance and self-reported smoking.

<table>
<thead>
<tr>
<th>Health surveillance</th>
<th>Non-smokers</th>
<th>Smokers</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Detection threshold = 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>3160</td>
<td>93</td>
<td>3253</td>
</tr>
<tr>
<td></td>
<td>(82.38)</td>
<td>(36.05)</td>
<td>(79.46)</td>
</tr>
<tr>
<td>Smokers</td>
<td>676</td>
<td>165</td>
<td>841</td>
</tr>
<tr>
<td></td>
<td>(17.62)</td>
<td>(63.95)</td>
<td>(20.54)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>3836</td>
<td>258</td>
<td>4094</td>
</tr>
<tr>
<td><strong>Detection threshold = 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>3548</td>
<td>144</td>
<td>3692</td>
</tr>
<tr>
<td></td>
<td>(92.49)</td>
<td>(55.81)</td>
<td>(30.18)</td>
</tr>
<tr>
<td>Smokers</td>
<td>288</td>
<td>114</td>
<td>402</td>
</tr>
<tr>
<td></td>
<td>(7.51)</td>
<td>(44.19)</td>
<td>(9.82)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>3836</td>
<td>258</td>
<td>4094</td>
</tr>
<tr>
<td><strong>Detection threshold = 3</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>3668</td>
<td>175</td>
<td>3843</td>
</tr>
<tr>
<td></td>
<td>(95.62)</td>
<td>(67.83)</td>
<td>(93.87)</td>
</tr>
<tr>
<td>Smokers</td>
<td>168</td>
<td>83</td>
<td>251</td>
</tr>
<tr>
<td></td>
<td>(4.38)</td>
<td>(32.17)</td>
<td>(6.13)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>3836</td>
<td>258</td>
<td>4094</td>
</tr>
<tr>
<td><strong>Detection threshold = 4</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>3736</td>
<td>191</td>
<td>3927</td>
</tr>
<tr>
<td></td>
<td>(97.39)</td>
<td>(74.03)</td>
<td>(95.92)</td>
</tr>
<tr>
<td>Smokers</td>
<td>100</td>
<td>67</td>
<td>167</td>
</tr>
<tr>
<td></td>
<td>(2.61)</td>
<td>(25.97)</td>
<td>(4.08)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>3836</td>
<td>258</td>
<td>4094</td>
</tr>
</tbody>
</table>

Note: Detection threshold is defined as the minimal number of positive peer reports that is required to classify a student as a smoker. If a student receives more than or equal to the required number of peer reports, the student will be classified as a smoker. Column percent in parentheses.

Having examined the smoking detection networks, we show how peer detections can be used to triangulate self-reports of smoking behavior. When using peer detections to cross-validate self-reports, we take self-reported smoking as a given and focus on correcting nonsmoking status, as students tend to under-report their smoking status (Kenkel et al., 2003). Table 3 includes a detailed breakdown of the distribution of students detected as smokers by their self-reported smoking status. We find about 18, 8, 4, and 3 percent of the students who self-identify as non-smokers are detected for smoking by at least one, two, three, and four peers, respectively. These students may have under-reported their smoking.

The question then becomes what is a proper detection threshold (i.e., the minimal number of positive peer reports) to classify a self-reported non-smoker as a smoker. Using one positive peer report as the threshold obviously yields the highest smoking rate. But it may create excessive false positives. Using too large a number of positive peer reports as the threshold, on the other hand, may be too conservative to provide sufficient correction to the under-reports. Three positive peer reports represent the 3rd quartile of the number of positive peer reports across the schools excluding zeros. Using three and four positive peer reports as thresholds result in between 100 and 168 self-reported non-smokers being classified as smokers or an increase in the total number of smokers by between 39 and 65 percent. Hence, even using relatively high thresholds, like three or four positive peer reports, uncovers a significant number of smokers that self-reports fail to reveal.

4.7.1. Sensitivity analysis

However, triangulations as such may still create false positives, namely, mistakenly identifying a nonsmoker as a smoker. To gain an understanding of the possible range of the false positives, we perform sensitivity analysis. We assume peer reports are independent from one another. Suppose the reporting accuracy \( P(X_i = 0|Y = 0) \) is 0.7, where \( Y \) represents a student’s actual smoking status and \( X_i \) a peer report. Then, the probability of a false positive created by peer reporting is \( P(X_i = 1|Y = 0) = 0.3 \). With three positive peer reports, the probability of the overall false positive (i.e., falsely identify a non-smoker as a smoker) is \( 0.3^3 = 0.027 \). Similarly, with four positive peer reports, the probability of the overall false positive becomes even smaller, \( 0.3^4 = 0.0081 \). Both are under the conventional 5% significance level. Of course, the probability of false positives depend on our assumption of the reporting accuracy. In Table 4 we show the results under a wide range of reporting accuracy. The results indicate that with a reasonable degree of reporting accuracy, three to four positive peer reports are good enough to reduce the probability of false detections.

4.7.2. Empirical benchmarking

Table 5 shows the resulting proportion of smokers in each school using a range of different thresholds, in which we classify a non-smoker as a smoker if the number of positive peer reports the student receives is larger or equal to the thresholds. Because the schools are from the same site and share similar profiles, the true smoking rates in these schools are likely similar. Thus, the optimal detection threshold should produce smoking rates that are least variable across the schools. If we use a uniform detection threshold, then Table 5 suggests that among the reasonable and comparable alternatives, using three or four positive peer reports leads to smaller variations in the refined smoking rates across the schools (SD = 0.0299 and 0.0319, respectively). If we allow the detection threshold to vary by school, then Table 5 suggests that the 3rd quartile of the received smoking detections in each school can be a good detection threshold, which appears to have produced the smallest variation in the refined smoking rates (SD = 0.0243). This is to be expected as each school may have its own optimal detection threshold because of its particular local culture.

According to the latest China national survey (Chinese Center for Disease Control and Prevention 2014), part of the Global Youth Tobacco Survey (GYTS), the proportion of current smokers (i.e., those smoked in the past 30 days) among middle school students in the site of our study is 7.5%. This number is the average smoking rate of the middle school students, of which 62% are from rural schools and the rest are from urban schools. Also note that the smoking rate of rural middle school students are 2.8% higher on average than that of urban middle school students. Thus, let \( R \) and \( U \) denote the smoking rate among rural and urban students, respectively. We have the following:

\[
7.5% = 62\% \times R + 38\% \times U, \quad \text{where} \quad U = R - 2.8%.
\]

Solving for \( R \), the actual smoking rate in our sample (i.e., rural middle school students by the definition of GYTS) should be about 8.7%, without accounting for possible under-reports. Assuming a

\[\text{Note: conventional sensitivity and specificity analysis (Benowitz et al., 2009; Kim and Jung, 2013; Wagenerknecht et al., 1992) does not work well in our case. First, it is because the analysis inappropriately assumes that self-reports are the golden standard, and uses them to calculate the sensitivity and specificity. Second, it is because sensitivity, specificity, and their sums are strictly linear functions of the detection thresholds in our case. See Table S3.}\]
Table 4
Sensitivity analysis of the overall false positive probability.

|                      | Reporting accuracy $P(X = 0|Y = 0)$ | False-positive probability $P(X = 1|Y = 0)$ | Overall false positive probability | Overall false positive probability | Overall false positive probability | Overall false positive probability |
|----------------------|--------------------------------------|--------------------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|
| 2 positive peer reports | 0.60                                 | 0.40                                       | 0.16                              | 0.06                              | 0.03                              | 0.01                              |
| 3 positive peer reports | 0.60                                 | 0.30                                       | 0.16                              | 0.09                              | 0.03                              | 0.01                              |
| 4 positive peer reports | 0.60                                 | 0.20                                       | 0.16                              | 0.09                              | 0.03                              | 0.01                              |
| 5 positive peer reports | 0.60                                 | 0.20                                       | 0.16                              | 0.09                              | 0.03                              | 0.01                              |
| 6 positive peer reports | 0.60                                 | 0.10                                       | 0.16                              | 0.09                              | 0.03                              | 0.01                              |

Note: Even when the reporting accuracy is only marginally better than flipping a coin where $P(X = 0|Y = 0) = 0.5$, the overall false positive probability is around or well below the 5% significance level if a student receives three or more positive peer reports.

Table 5
Proportion of smokers based on a refined measure of smoking combining peer detection and self-reports.

<table>
<thead>
<tr>
<th>Self-reported Detection threshold</th>
<th>School 1</th>
<th>School 2</th>
<th>School 3</th>
<th>School 4</th>
<th>School 5</th>
<th>School 6</th>
<th>SD</th>
<th>Overall (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.04</td>
<td>0.12</td>
<td>0.06</td>
<td>0.06</td>
<td>0.07</td>
<td>0.05</td>
<td>0.0280</td>
<td>6.30</td>
</tr>
<tr>
<td>Median</td>
<td>0.18</td>
<td>0.28</td>
<td>0.24</td>
<td>0.17</td>
<td>0.23</td>
<td>0.28</td>
<td>0.0473</td>
<td>22.81</td>
</tr>
<tr>
<td>3rd Quartile</td>
<td>0.08</td>
<td>0.15</td>
<td>0.08</td>
<td>0.07</td>
<td>0.09</td>
<td>0.07</td>
<td>0.0322</td>
<td>7.79</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(1)</td>
<td>(1)</td>
<td>(2)</td>
<td>(1)</td>
<td></td>
<td>0.0243</td>
<td>11.60</td>
</tr>
</tbody>
</table>

Note: Detection threshold is defined as the minimal number of positive peer reports that is required to classify a nonsmoker as a smoker. Median and 3rd Quartile are the median and 3rd quartile of the number of received detections excluding zero in each school. They are used as additional thresholds and vary across schools. The corresponding detection thresholds are in parentheses below the proportions for each school. SD is the standard deviation of the smoking rates across schools. Overall is the overall smoking rate among all students across schools.

A reasonable degree of under-reporting, for example, between 15 and 30% of the smokers under-reported their smoking, then the actual smoking rate should be around the range (9.8%, 11.4%). In Table 5, we can see that using three positive peer reports to correct under-reports leads to a final smoking rate that is right in this range while using four positive peer reports and the 3rd quartile of the received smoking detections in each school as detection thresholds lead to final smoking rates that are next closest to the projected range. In contrast, using five positive peer reports produces too low a final smoking rate while using one or two positive peer reports leads to too high a final smoking rate.

In short, our analysis shows that using three or four positive peer reports as detection thresholds helps to uncover a significant number of under-reported smokers while unlikely producing excessive false positives. However, to what extent this finding is generalizable to other contexts is an empirical question. The optimal thresholds may vary by the specific contexts and research goals. The overall recommendation we make here is a probabilistic one: the more positive peer reports an individual gets, the more likely the individual has a particular health behavior.

5. Conclusion and discussion

In this study, we propose a network method for health surveillance and use a novel dataset on smoking detection networks to demonstrate the method. We find that smokers, optimistic students, and popular students make better informants in terms of detecting peers smoking. On one hand, this confirms the conventional wisdom in related methods like respondent-driven sampling of using subjects with the characteristic of interest as seeds to approach hard-to-reach subjects. On the other hand, it points out that best informants are not exclusively or necessarily subjects with the characteristic of interest. Other characteristics like popularity and personality also matters for being effective informants. We also find smoking detection tends to be homophilous in socioeconomic status – students from similar socioeconomic background are more aware of one another’s smoking than those from different backgrounds. Smoking detection also seems to align with the direction of friendships (students are more aware of the smoking of their nominated friends while the converse is not true).

We also show that peer reports can be used to address self-reporting bias. We demonstrate that using three to four (or the third quartile of) positive reports helps to uncover a good number of under-reported smokers while unlikely producing excessive false positives. To what degree our finding is applicable to other situations is worth further investigations. But the methods we used, including both the sensitivity analysis and the empirical benchmarking, may be informative for future similar explorations. Also, the general principle we advocate – the more positive reports an individual gets, the more likely the individual has a certain health behavior – should be quite robust across culture and context.

On the surface, our health surveillance method appears to be similar to chain-referral methods like respondent-driven sampling (RDS) (Salganik and Heckathorn, 2004; Gile and Handcock, 2010). In the RDS, a number of seed nodes are used to recruit other similar nodes and the latter group is further used to recruit more similar nodes. The main purpose of RDS is to use these chains of reports to estimate the size of “hidden populations”. In contrast, the main
purposes of our method are to identify good informant for monitoring subsequent health behaviors and to triangulate self-reports. However, connections can indeed be drawn between the two methods. For example, the informants identified by our method may be used as the initial seed nodes in the RDS. By properly addressing self-reporting bias, our health surveillance can also provide an accurate estimate of the prevalence of health behaviors. In Table S4 we list more detailed comparisons between the two methods.

Researchers may be concerned with the practicality of our method and its feasibility compared to biological tests. In Table S5 we provide detailed comparisons between the two methods. In short, our method is more convenient and affordable to use (as it does not require specialized instruments and professional staff to implement), and is less intrusive to subjects. However, health surveillance and biological tests are not meant to be competing methods. They can be used in tandem, depending on the researcher’s interests and resources.

Several limitations in our data or analysis are worth pointing out. First, there are perhaps strong norms against students reporting on their peers’ deviant behaviors like smoking (Wilson-Simmons et al., 2006), because we found that many students did not report any smokers at all. Have the students been more active in reporting on their peers, more accurate triangulation may be possible.

Second, we asked each student to report no more than four other students in the smoking detection networks. This limits the information we can extract from these reports. It is possible that allowing more nominations will lead to more complete information about student smoking. Thus the results presented in this paper are a conservative estimation of the feasibility and effectiveness of the health surveillance method. We advise future research not to place a cap on peer detections. This would greatly enhance the feasibility and generalizability of the health surveillance method.

Third, we are unable to determine the accuracy of peer reports, but gaining a better understanding of peers reporting accuracy helps to improve triangulations of self-reports. One potential method for dealing with this issue is Butts’ approach to combining multiple peer reports (Butts, 2003; Marcum et al. 2012).

However, in our study, the focus is on refining measurement of the attributes (i.e., smoking status) of the actors rather than the connections between them. In addition, our detection networks are quite sparse and so may not provide sufficient information to appropriately identify each student’s reporting accuracy. Thus Butts’ method, despite its great potential, is not directly applicable in our case. In addition, we may weight peer reports differently according to their network positions, as previous and our research has found that popular actors usually make more accurate reports (e.g., Romney and Faust, 1982; Freeman et al., 1987; Freeman, 1992; Krackhardt and Kilduff, 1999).

Last, the data we used comes from a particular site in China. To what extent our findings are generalizable to other schools and countries is unclear. More research in this line would be useful to cross-validate or extend our findings. For example, our method may be used to monitor binge-drinking in college or health behaviors in retirement communities, etc. In sum, monitoring health behaviors through social networks is an exciting new area of research with many potential applications, opportunities, and challenges.

Acknowledgements

We thank Christopher Winship, Martin K. Whyte, James O’Malley, Stephen Benard, Jessica Calarco, Youngjoo Cha, Jennifer C.

Lee, and Cate Taylor for their suggestions in the early development of this project. The Graduate School of Arts and Sciences, the Multidisciplinary Program in Inequality and Social Policy, the Fairbank Center for Chinese Studies, and the Institute for Quantitative Social Science, all at Harvard University, provided financial support for this project. The National Science Foundation also provided financial support through a Graduate Research Fellowship (NSF Grant DGE-0813962) to the second author.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.socnet.2015.02.001.

References


10 Additional analyses [see Table S6] show that artificially restricting the nominations in the smoking detection networks (and the cigarette exchange networks) to smaller numbers does not significantly alter our main results.


