Social Networks 42 (2015) 8–17

Contents lists available at ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

Health surveillance through social networks

Weihua An [a,*], Long Doan [b]

[a] Indiana University, Department of Sociology and Department of Statistics, 752 Ballantine Hall, 1020 E. Kirkwood Ave., Bloomington, IN 47405, USA
[b] Indiana University, Department of Sociology, 744 Ballantine Hall, 1020 E. Kirkwood Ave., Bloomington, IN 47405, USA

ARTICLE INFO

Keywords:
Health surveillance
Social networks
Informants
Reporting bias
Peer reports
Smoking

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.

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.

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; Cornwell, 2009; Cornwell 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: weihuaan@indiana.edu (W. An), longdoan@indiana.edu (L. Doan).

http://dx.doi.org/10.1016/j.socnet.2015.02.001
0378-8733/

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 peer 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.