Original Paper

Tobacco Use Behaviors, Attitudes, and Demographic Characteristics of Tobacco Opinion Leaders and Their Followers: Twitter Analysis

Kar-Hai Chu[1*], PhD; Anuja Majmundar[2*], MBA, MA; Jon-Patrick Allem[2*], PhD; Daniel W Soto[2*], MPH; Tess Boley Cruz[2], PhD; Jennifer B Unger[2*], PhD

[1]University of Pittsburgh, Pittsburgh, PA, United States

[2]University of Southern California, Los Angeles, CA, United States

[*]these authors contributed equally

Corresponding Author:
Kar-Hai Chu, PhD
University of Pittsburgh
230 McKee Place, Suite 600
Pittsburgh, PA,
United States
Phone: 1 412 692 2578
Email: chuk@pitt.edu

Abstract

Background: Tobacco-related content on social media is generated and propagated by opinion leaders on the Web who disseminate messages to others in their network, including *followers*, who then continue to spread the information. Opinion leaders can exert powerful influences on their followers' knowledge, attitudes, and behaviors; yet, little is known about the demographic characteristics and tobacco use behavior of tobacco opinion leaders on the Web and their followers, compared with general Twitter users.

Objective: In this study, we hypothesized that opinion leaders use more tobacco products and have higher nicotine dependence than the other 2 groups (eg, followers and general Twitter users) and that followers—those who spread messages by opinion leaders—would more likely be in demographic groups that are vulnerable to tobacco marketing influence (eg, young adults and lower educational attainment).

Methods: We constructed the social networks of people who tweet about tobacco and categorized them using a combination of social network and Twitter metrics. To understand the characteristics of tobacco opinion leaders and their followers, we conducted a survey of tobacco opinion leaders, their followers, and general Twitter users. The sample included 347 opinion leaders, 567 followers, and 519 general users. The opinion leaders had a median of 1000 followers, whereas followers and general users had fewer than 600 followers.

Results: Opinion leaders were more likely than their followers to report past month use of tobacco products; followers, in turn, were more likely to report past month use of these products than general Twitter users. The followers appeared to be an especially vulnerable group; they tended to be younger (mean age 22.4 years) and have lower education compared with the opinion leaders and general users.

Conclusions: Followers of Twitter tobacco opinion leaders are a vulnerable group that might benefit from antitobacco education to counter the protobacco communications they see on social media.

XSL•FO
RenderX

Introduction

Social media has emerged as a popular forum for tobacco product users and prospective new users to learn about and discuss nicotine and tobacco products and for businesses to promote these products. Previously identified themes of tobacco-related social media posts include marketing and promotions by manufacturers, posts by tobacco users about their own tobacco experiences, discussions about flavors, and debates about tobacco product regulations [1-3]. Tobacco-related social media posts also contain content that could entice youth and nontobacco users to try tobacco products, including cartoons and other youth-oriented themes [4,5], hookah paired with alcohol in social settings [6], little cigar or cigarillos paired with marijuana [7], and pictures of people blowing large clouds of e-cigarette aerosol [8,9].

As social media messages about tobacco products can influence perceptions about the health effects and potential harms of these products and the social norms of use [10], it is important to understand as much as possible about these messages and the people who are disseminating them. Although previous research has focused on the content of tobacco-related messages on social media, less is known about the people who generate, receive, and propagate those messages. Previous analyses of information flow in Twitter have shown that a small number of *elite* users (approximately 20,000 people) generate nearly half of the tweets [11]. Only about 15.02% (1377/9165) of tweets received by ordinary users are directly from traditional mass media sources (eg, Cable News Network); most are filtered through opinion leaders who selectively retweet information from mass media sources [11]. Twitter opinion leaders—people who occupy central positions in their social networks, have numerous well-connected followers, have social status and credibility, and are emulated by their followers—are important members of Web-based communities as they control which information diffuses through social networks [12].

Opinion leaders discussing tobacco products on Twitter can potentially influence their followers to try new products, adopt beliefs about the relative harm of tobacco products, and support or oppose tobacco control policies. Thus, it is important to understand who these tobacco-related opinion leaders are and how their personal attitudes and behaviors might be influencing the discourse on the Web about tobacco products.

Intervention or education programs can benefit by leveraging opinion leaders to champion their ideas [13].

This study identified tobacco-related opinion leaders on Twitter by combining Twitter user metadata with techniques in social network analysis to develop a more comprehensive definition of opinion leaders. We then conducted surveys of these opinion leaders, people who follow these opinion leaders, and general Twitter users who are not engaged in tobacco-related discussions. We compared social network characteristics, demographic characteristics, tobacco product use, and nicotine dependence to identify differences among opinion leaders, followers, and general Twitter users. Opinion leaders, in this research, are operationalized as individuals strategically situated in their social networks. Their messages are disseminated widely via *shares* or *retweets* as they are viewed as subject matter experts. Research suggests that such high involvement and engagement in specific topics leads individuals to raise awareness about those topics and transition to polytobacco product use [14,15]. In keeping with this behavior, we hypothesized that opinion leaders would use more tobacco products and have higher nicotine dependence than the other 2 groups. We also hypothesized that followers, operationalized as individuals who predominantly follow and disseminate social media messages of opinion leaders, would more likely be in demographic groups that are vulnerable to the influence of tobacco marketing (eg, young adults and racial or ethnic minorities).

Methods

Data Collection

Twitter data were obtained with a custom Java 7 program based on Twitter4J v.4.0.3 that continuously accesses the Twitter streaming application programming interface (API) v.1.1 and collects tweets that contain any 1 of over 200 tobacco-related keywords, for example, cigarette, e-cigarette, or vape (see Multimedia Appendix 1). Twitter data were collected from March 2015 to March 2016. Along with the text of the tweet, the data include Twitter metadata such as the user name of the person who posted the tweet and whether the tweet was an original tweet or a retweet. This information was used to construct the retweet network by retrieving the data of every user who posted tobacco-related content.

Solely using Twitter metrics to identify opinion leaders can be misleading. Nontraditional accounts (eg, celebrities) can distort actual user classifications, whereas the number of followers is more likely to measure popularity rather than influence [16,17]. Therefore, we applied a combination of social network analysis–clustering algorithms and Twitter metrics to classify 3 types of individuals as follows: an opinion leader, a follower, and a general Twitter user. Opinion leaders and followers are additionally defined as users in our data who had posted tobacco-related content compared with general Twitter users who did not. First, a network was generated by linking users who had retweeted another user; this resulted in a retweet network defined by ties between the person who posted a tweet and the person who retweeted it. From this network, clusters were identified by conducting a modularity analysis. Modularity helps identify clusters within a network by grouping nodes (ie, Twitter users) that have more connections (ie, retweets) with others within a group than those outside of the group [18]. After the clusters were identified, opinion leaders were chosen as those who had been retweeted the most; followers were identified within each cluster as those who had retweeted others the most. Independently, general Twitter users were found by the Twitter API's get-user-status function, which returns users who have recently posted a tweet about any topic (not just tobacco).

This method produced a convenience sample of 347 opinion leaders, 567 followers, and 519 general users. We sent Twitter private messages to potential participants inviting them to complete the survey. Each private message contained a unique,

randomly generated link to a RedCap site where the survey was hosted. Clicking on the link identified the respondent as an opinion leader, follower, or general user who had been invited to complete the survey. This was done so that only people who received an invitation link could complete the survey. When respondents clicked on the link and arrived on the RedCap survey page, they saw an institutional review board–approved consent script. After clicking on a button indicating their consent to participate, they were directed to the survey. Participants received a US $20 gift card for completing the survey. Networks were constructed in April 2016. Surveys were sent out from May 2016 to June 2018.

Measures

Participants self-reported their age, sex, race and ethnicity, and education. Social network characteristics were assessed by asking participants how many Twitter users they followed and how many Twitter users followed them. The survey asked which of the following products the participants had used in the past month: cigarettes, e-cigarettes, cigars, pipe tobacco, blunts, hookah, smokeless tobacco, cigarillo, marijuana, and alcohol.

Statistical Analysis

The 3 groups (opinion leaders, followers, and general users) were compared on all measures, using analysis of variance for normally distributed continuous variables, the Kruskal-Wallis test for nonnormally distributed continuous variables, and chi-square for categorical variables.

Results

Demographic Differences Across Groups

The sample included 1433 completed surveys—347 opinion leaders, 567 followers, and 519 general users. The followers (mean age 22.4 years) were significantly younger than the opinion leaders (mean age 24.2 years) and the general users (mean age 25.2 years), $P<.001$. Compared with opinion leaders and general users, followers were more likely to be Hispanic ($P=.03$). General users were more likely than opinion leaders and followers to be African American ($P=.03$). Compared with opinion leaders and general users, followers had less education: only 11% (40/380) of followers had a bachelor's degree or higher as compared with 19% (45/241) of opinion leaders and 26% (92/350) of general users ($P<.001$).

Tobacco or Nicotine Product and Other Substance Use

For most of the tobacco products, opinion leaders reported the highest past month use prevalence, followed by followers and general users (Table 1). This pattern was evident for cigarettes, e-cigarettes, cigars, blunts, hookah, and cigarillos. Opinion leaders had the highest nicotine dependence scores, followed by followers and general users. Opinion leaders were also more likely than followers and general users to have used alcohol or marijuana in the past month.

Table 1. Comparison of opinion leaders, followers, and general users.

Twitter user characteristics	Opinion leaders (n=347)	Followers (n=567)	General users (n=519)	*P* value
Social network size				
Number of Twitter users who follow the respondent (median)	1000	554	503	.001
Number of Twitter users whom the respondent follows (median)	428	375	366	.01
Age (years)	24.2	22.4	25.2	<.001
Female, n/N (%)	122/242 (50.4)	211/392 (53.8)	202/360 (56.1)	.39
Race and ethnicity, n/N (%)				
African American	26/242 (11)	40/392 (10)	67/360 (19)	.03
Asian or Pacific Islander	15/242 (6)	21/392 (5)	14/360 (4)	.03
Hispanic	51/242(21)	102/392 (26)	74/360 (21)	.03
White	129/242 (53)	201/392 (51)	176/360 (49)	.03
Other	21/242 (9)	28/392 (7)	29/360 (8)	.03
Education, n/N (%)				
High school or less	85/241 (35)	155/380 (41)	109/350 (31)	<.001
Some college	111/241 (46)	185/380 (49)	149/350 (43)	<.001
Bachelor's degree or higher	45/241 (19)	40/380 (11)	92/350 (26)	<.001
Number of tobacco products used in past month (mean)	1.31	.98	.66	<.001
Use of specific products in the past month, n/N (%)				
Cigarettes	93/347 (27)	117/567 (21)	83/519 (16)	.01
E-cigarettes	65/347(19)	93/567 (16)	50/519 (10)	.01
Cigar	30/34 (9)	37/567 (7)	23/519 (4)	.04
Pipe	37/347 (11)	64/567 (11)	28/519 (6)	.01
Blunt	96/347 (28)	118/567 (21)	73/519 (14)	.01
Hookah	45/347 (13)	51/567 (9)	29/519 (6)	.01
Smokeless	13/347 (4)	7/567 (1)	11/519 (2)	.04
Cigarillo	60/347 (17)	67/567 (12)	41/519 (8)	.01
Other substance use in the past month, n/N (%)				
Alcohol	159/229 (69)	221/378 (58)	215/354 (61)	.02
Marijuana	103/219 (47)	142/371 (38)	96/330 (29)	<.001
Nicotine dependence score (mean)	1.48	1.18	1.02	.04
Number of tobacco brands followed on Twitter (mean)	.23	.18	.13	.11

Discussion

Principal Findings

Findings suggest that opinion leaders were more likely to report past month use of tobacco products than their followers; followers, in turn, were more likely to report past month use of these products than general Twitter users. The followers appeared to be an especially vulnerable group; they tended to be younger and have lower education. Opinion leaders had higher nicotine dependence scores and were more likely to report past month alcohol or marijuana use compared with followers and general users.

Tobacco opinion leaders on Twitter use a wide variety of tobacco products and other substances. As opinion leaders are typically held in high esteem by their followers, they play an important role in establishing and conveying social norms [19]. Opinion leaders who discuss their polytobacco and polysubstance use on Twitter might lead their followers to believe that these behaviors are normative, safe, or socially admirable. Followers, in turn, might emulate opinion leaders' levels of tobacco use and become nicotine dependent.

Social media–based tobacco campaigns can address tobacco use disparities by tailoring messages that resonate with followers. Such focused efforts can potentially play an important role in educating followers who are typically younger and less

educated than the other groups. Past evidence suggests that network-based interventions that involve identifying peer messengers result in improved health behaviors and more targeted delivery of interventions [13,20]. Social network analysis of a social media–based intervention also revealed that participants from vulnerable demographic groups (younger youth and females) may require additional outreach efforts [21]. Future tobacco health communication campaigns can take advantage of strategic delivery of health messages to followers on social media.

Solely using Twitter metrics to identify opinion leaders can be misleading, as bots, celebrities, and other nontraditional accounts can distort actual user classifications; Twitter metrics such as the number of followers are more likely to measure popularity rather than influence [16,17]. By using social network analysis in combination with Twitter metrics in this study, we are able to systematically identify emergent clusters in the Twitter tobacco network and then apply Twitter metrics to identify subgroup opinion leaders. This method helps prevent over-reliance on Twitter metrics such as follower or retweet count as the sole metric to define opinion leadership.

Limitations

The study utilized a convenience sample by sending unsolicited messages to Twitter users. The tobacco opinion leaders, followers, and general Twitter users were selected on the basis of their positions in the Twitter social network; they had not previously expressed interest in participating in surveys. Twitter users who read their direct messages, click on a survey link, and complete a Web-based survey might not be representative of the general Twitter population; in addition, we have no method to verify that the user who takes the survey is the same as the original Twitter user who received the link.

Conclusions

Despite these limitations, these findings provide important new information about people who disseminate and receive tobacco-related information on Twitter. Opinion leaders are influential as they occupy central positions in the social network and have the potential to communicate with a wide audience of Twitter users. Our findings indicate that tobacco opinion leaders use a wide variety of tobacco products as well as other substances. They may disseminate these attitudes to their Twitter followers who tend to be members of vulnerable populations (eg, young adults and lower educational attainment). Over time, repeated exposure to messages from tobacco opinion leaders could place followers at an increased risk for tobacco product experimentation and escalation. Although this survey was restricted to Twitter users aged 18 years and older, it is likely that younger Twitter users also follow tobacco opinion leaders, and these opinion leaders' messages could persuade them to experiment with tobacco. This study demonstrates that it is possible to identify tobacco opinion leaders on Twitter and their followers and opens the opportunity to apply other methods of supplementing Twitter measures to classify Twitter users. Opinion leaders on the Web have large, well-connected social networks of social media users who may look to them for information, opinions, and advice. If the information disseminated by opinion leaders on the Web is incorrect or biased, their followers could make important decisions based on faulty information. Future research should determine how opinion leaders influence their followers' offline tobacco behaviors.

Authors' Contributions

All authors contributed to writing the manuscript and approved the final draft. KC developed the algorithm for constructing the Twitter social networks and identifying opinion leaders, followers, and general Twitter posters. AM conducted data analysis and contributed to the interpretation of findings. DS contributed to conceptualization of the study, oversaw data collection, and contributed to the interpretation of the findings. JPA contributed to conceptualization of the study and interpretation of findings. TBC contributed to conceptualization of the study and interpretation of findings. JU contributed to conceptualization of the study, data analysis, and interpretation of findings.

Conflicts of Interest

None declared.

Multimedia Appendix 1

List of tobacco-related keywords.

[DOCX File, 32KB-Multimedia Appendix 1]

References

1. Ayers JW, Leas EC, Allem J, Benton A, Dredze M, Althouse BM, et al. Why do people use electronic nicotine delivery systems (electronic cigarettes)? A content analysis of Twitter, 2012-2015. PLoS One 2017;12(3):e0170702 [FREE Full text] [doi: 10.1371/journal.pone.0170702] [Medline: 28248987]

2. Cole-Lewis H, Pugatch J, Sanders A, Varghese A, Posada S, Yun C, et al. Social listening: a content analysis of e-cigarette discussions on Twitter. J Med Internet Res 2015;17(10):e243 [FREE Full text] [doi: 10.2196/jmir.4969] [Medline: 26508089]

3. Zhan Y, Liu R, Li Q, Leischow SJ, Zeng DD. Identifying topics for e-cigarette user-generated contents: a case study from multiple social media platforms. J Med Internet Res 2017 Jan 20;19(1):e24 [FREE Full text] [doi: 10.2196/jmir.5780] [Medline: 28108428]

4. Kirkpatrick MG, Cruz TB, Goldenson NI, Allem J, Chu K, Pentz MA, et al. Electronic cigarette retailers use Pokémon Go to market products. Tob Control 2017 Dec;26(e2):e145-e147. [doi: 10.1136/tobaccocontrol-2016-053369] [Medline: 28044010]

5. Allem J, Cruz TB, Unger JB, Toruno R, Herrera J, Kirkpatrick MG. Return of cartoon to market e-cigarette-related products. Tob Control 2018 Jul 26:-. [doi: 10.1136/tobaccocontrol-2018-054437] [Medline: 30049688]

6. Allem J, Chu K, Cruz TB, Unger JB. Waterpipe promotion and use on Instagram: #hookah. Nicotine Tob Res 2017 Jan 11;19(10):1248-1252. [doi: 10.1093/ntr/ntw329] [Medline: 28077449]

7. Kostygina G, Tran H, Shi Y, Kim Y, Emery S. 'Sweeter Than a Swisher': amount and themes of little cigar and cigarillo content on Twitter. Tob Control 2016 Oct;25(Suppl 1):i75-i82 [FREE Full text] [doi: 10.1136/tobaccocontrol-2016-053094] [Medline: 27697951]

8. Chu K, Allem J, Cruz TB, Unger JB. Vaping on Instagram: cloud chasing, hand checks and product placement. Tob Control 2016 Sep 22:575-578. [doi: 10.1136/tobaccocontrol-2016-053052] [Medline: 27660111]

9. Unger JB, Escobedo P, Allem J, Soto DW, Chu K, Cruz T. Perceptions of secondhand e-cigarette aerosol among Twitter users. Tob Regul Sci 2016 Apr;2(2):146-152 [FREE Full text] [doi: 10.18001/TRS.2.2.5] [Medline: 28090560]

10. Depue JB, Southwell BG, Betzner AE, Walsh BM. Encoded exposure to tobacco use in social media predicts subsequent smoking behavior. Am J Health Promot 2015;29(4):259-261. [doi: 10.4278/ajhp.130214-ARB-69] [Medline: 24670071]

11. Wu S, Hofman J, Mason W, Watts D. Who says what to whom on Twitter. 2011 Presented at: International World Wide Web Conference; March 28-April 1, 2011; Hyderabad, India p. 705-714 URL: http://www.wwwconference.org/proceedings/www2011/proceedings/p705.pdf

12. Park CS, Kaye BK. The tweet goes on: interconnection of Twitter opinion leadership, network size, and civic engagement. Comput Hum Behav 2017 Apr;69:174-180. [doi: 10.1016/j.chb.2016.12.021]

13. Valente TW. Network interventions. Science 2012 Jul 6;337(6090):49-53. [doi: 10.1126/science.1217330] [Medline: 22767921]

14. Feick L, Price LL. The market maven: a diffuser of information marketplace. J Market 1987 Jan;51(1):83-97. [doi: 10.2307/1251146]

15. Soneji S, Yang J, Knutzen KE, Moran MB, Tan AS, Sargent J, et al. Online tobacco marketing and subsequent tobacco use. Pediatrics 2018 Jan 2;141(2):e20172927. [doi: 10.1542/peds.2017-2927]

16. Riquelme F, González-Cantergiani P. Measuring user influence on Twitter: a survey. Inf Process Manag 2016 Sep;52(5):949-975. [doi: 10.1016/J.IPM.2016.04.003]

17. Garcia D, Mavrodiev P, Casati D, Schweitzer F. Understanding popularity, reputation, and social influence in the Twitter society. Policy Intern 2017;9(3):64. [doi: 10.1002/poi3.151]

18. Newman MEJ. Modularity and community structure in networks. Proc Natl Acad Sci U S A 2006 Jun 6;103(23):8577-8582 [FREE Full text] [doi: 10.1073/pnas.0601602103] [Medline: 16723398]

19. Weeks BE, Ardèvol-Abreu A, Gil de Zúñiga H. Online influence? Social media use, opinion leadership, and political persuasion. Int J Public Opin Res 2015 Dec 31:edv050. [doi: 10.1093/ijpor/edv050]

20. Rice E, Yoshioka-Maxwell A, Petering R, Onasch-Vera L, Craddock J, Tambe M, et al. Piloting the use of artificial intelligence to enhance HIV prevention interventions for youth experiencing homelessness. J Soc Social Work Res 2018 Dec;9(4):551-573. [doi: 10.1086/701439]

21. Rice E, Tulbert E, Cederbaum J, Barman Adhikari A, Milburn NG. Mobilizing homeless youth for HIV prevention: a social network analysis of the acceptability of a face-to-face and online social networking intervention. Health Educ Res 2012 Jan 13;27(2):226-236. [doi: 10.1093/her/cyr113]

Abbreviations

API: application programming interface
FDA: Food and Drug Administration
NIH: National Institutes of Health

XSL•FO
RenderX

XSL•FO

RenderX