Jumpstarting Online Interactions to Promote Smoking Cessation: Analyzing the Role of Seed Users in an Online Health Community

Xiangyu Wang[a], Kang Zhao[a], Michael S. Amato[bc], Cassandra A. Stanton[d], Jonathan Shuter[e] and Amanda L. Graham[bc]*

[a]The University of Iowa, Iowa City, IA. [b]Innovations Center, Truth Initiative, Washington, DC. [c]Department of Medicine, Mayo Clinic College of Medicine and Science, Rochester, MN. [d]Westat, Rockville, MD. [e]Department of Medicine, Albert Einstein College of Medicine, Bronx, NY.

Abstract

When starting a new online community, a useful practice is to recruit "seed users" to create content and encourage participation from community members. However, in the context of online health communities (OHCs), where users with similar health concerns interact, the impact of interacting with seed users on members' health outcomes remains unknown. Using data from an OHC for smoking cessation, we found that support from seed users predicts member abstinence. In addition, seed users who were former smokers were more effective in supporting others to quit. Text analytics also revealed differences between the support provided by seed users who were former smokers vs. current smokers. Outcomes of this study can aid the design of a supportive OHC and the promotion of smoking cessation.

Keywords: social support, seeding community, smoking cessation, text analytics.

Introduction

Online communities represent a popular means for people to interact. A successful online community relies on voluntary participation since most online communities do not offer any financial incentive for users to stay, and users are free to leave. A fledging online community can languish with little or even no activity (Young, 2013). One practice to promote online participation

is to "seed" content in new online communities by recruiting individuals to create content(Solomon & Wash, 2012).

An online health community (OHC) is an online community designed for a specific group of people with similar health concerns where people relate, exchange information, offer support and network with each other (McLeish & Redshaw, 2015). A sustainable and supportive online health community depends on members' active participation and the formation of meaningful long-term relationships (Iriberri & Leroy, 2009). For an OHC to take root and thrive, it is critical to build a core group of active members at its inception (Iriberri & Leroy, 2009). In many OHCs, seed users have been recruited to boost users' online engagement with duties such as starting discussions on certain topics, responding to new members' posts, establishing new connections, and encouraging lurkers to post (Resnick et al., 2010). Existing studies have examined different types of user behavior patterns and developed plans for maximizing communities' engagement at different stages (Bishop, 2007; Iriberri & Leroy, 2009; Stearns et al., 2014; Young, 2013). For example, lurkers feel more confident to join an online community when it is active and engaged (Bishop, 2007), and newcomers who have received a prompt support would encourage their continued participation (Xiangyu Wang et al., 2021; Young, 2013). Developing an active and engaged online community requires to provide the right tools, offer support and establish connections at the inception stage (Iriberri & Leroy, 2009). Strategic interventions, such as post recommendation or co-posting in the same thread, have been used to improve user engagement in OHCs (Stearns et al., 2014). However, it remains unknown if seed users can help others achieve desired health outcomes (e.g., smoking cessation) and if the choice of seed users should be made in a more strategic way.

One of the health outcomes we are particularly interested in is smoking cessation. Tobacco use is one of the most serious public health concerns in the U.S. (J. Baxter, 2017) and remains the leading preventable cause of premature death (Centers for Disease Control and Prevention, 2021). Especially among people infected with HIV, smoking is a major cause of morbidity and mortality. Tragically, many people survive their HIV/AIDS only to die from a tobacco-related illness. Previous work indicated that people living with HIV have a higher smoking prevalence (40%-80% vs. 20%) and are less likely to quit compared to the general population (Kim et al., 2017; Mdodo et al., 2015). The benefits of smoking cessation include improving health status, enhancing quality of life, increasing life expectancy and reducing the risk of disease (Centers for Disease Control and Prevention, n.d.). Quitting can be challenging; millions of tobacco users search online for cessation resources such as OHCs each year (Graham & Amato, 2019).

This study analyzed data from a seeded OHC that was created in the context of a randomized clinical trial for smoking cessation among people living with HIV/AIDS. Results suggest that online social support provided by seed users is a predictor of members' abstinence, with seed users who were ex-smokers being more effective in helping others to quit. Further analyses found that the language used by seed users who were ex-smokers to support others was linguistically different from support provided by seed users who were current smokers. The outcomes of this study can inform the design of a new OHC, guide the strategy to recruit seeders, and help OHC members quit smoking.

Related work

OHCs enable members to exchange social support in an anonymous environment with few limitations related to geography or time. Support received within an OHC can differ from and be

more powerful than support provided by family members. OHC members share similar health concerns that may not be well understood or appreciated by others without such identities (Lichtenstein et al., 1986). Research has found that social support from OHCs is associated with higher levels of online engagement (Xi Wang et al., 2017) and better psychological and physical outcomes (Cobb et al., 2010; Cohen, 1988; Wright et al., 2003). Specifically in OHCs for smoking cessation, those who received more social support from peers or had higher social network centralities were more likely to quit and less likely to relapse (Baskerville et al., 2015; Cheung et al., 2015; Graham et al., 2017; Myneni et al., 2015; Pechmann et al., 2017; Zhao et al., 2016). Previous work also indicated that different types of social support were associated with different exchange patterns: informational support is provided by users who are at late quit stages (i.e., established quitters with wisdom to share), while emotional support is exchanged among users who are at the same quit stage (i.e., current quitters experiencing the same struggles) (Ploderer et al., 2013; Zhang & Yang, 2015). Different types of social support can be associated with abstinence in different ways as well- informational support providers tend to have a longer duration of abstinence than emotional support providers, while informational support receivers have a shorter abstinent time than emotional support receivers (Zhang & Yang, 2015).

However, previous studies have mainly focused on social support exchanged within established communities among established members. In the early stages of an OHC, the number of active users and the amount of support they can provide to each other can be low and sporadic. Recruiting seed users is a common practice for an OHC to jumpstart online participation and support its early members. While the role of seed users in promoting online activities and engaging early users in an OHC has been well documented (Resnick et al., 2010; Young, 2013), the role that seed users

may play in promoting offline health behaviors is unclear. Therefore, this study attempts to answer two research questions specific to an OHC for smoking cessation:

RQ1: Is social support provided by seed users a predictor of OHC members' abstinence?

RQ2: Do seed users with different smoking statuses have different effects on the abstinence of OHC members they interacted with?

Answers to these questions can help OHC managers better understand and quantify the value of seed users to OHC members and inform the recruiting strategy of seed users.

Methods

Datasets

The data used in this study were collected from *Positively Smoke Free on the Web (PSFW)*, a web-based smoking cessation intervention with a fully integrated OHC created specifically for smokers living with HIV/AIDS. There are two types of users in this OHC: *Focal users* were HIV positive patients who smoked that were recruited to participate in a randomized clinical trial to evaluate the effectiveness of PSFW; *Seed users* were recruited as part of the study team and tasked with starting threads and providing support to focal users. A seed user could be an ex-smoker or a current smoker. The dataset spanned June 2016 to September 2020 and included 188 focal users and 7 seed users. Seed users did not have badges or other indicators to explicitly signal their special role in the community, however there were no restrictions on revealing those roles in the content of posts. For each focal user, data were available from a baseline survey (e.g., demographics) collected during study enrollment (prior to OHC use) and a follow-up survey that assessed smoking cessation outcomes 3 months after enrollment. Among the focal users, 18.08% (34 out of 188) reported 7-day point prevalence abstinence at 3 months. Among the seven seed users, four

quit before the start of the study and three were smoking at study inception. Back-end data were available for all focal and seed users' online interactions in the OHC.

Measures

The dependent variable in our logistic regression model was a binary variable that indicated whether a focal user reported 7-day abstinence at 3-month follow-up assessment. In this study, abstinence was defined by both a negative self-report ("No" to the question, "Have you smoked a cigarette, even a single puff, in the last 7 days?") and biochemical confirmation by exhaled carbon monoxide (ECO) . In this study, ECO<10ppm was considered to be in the non-smoker range, a standard methodology in tobacco treatment trials (Wagener et al., 2017). In other words, if a participant self-reported abstinence but had an ECO above 10ppm, his/her status was classified as "not abstinent."

Independent variables for our model measured focal users' online interactions with others from three different levels. For a focal user, an interaction could be one of her replies to others' post (a.k.a., "active interaction"), or a reply to his/her post (a.k.a., support-receiving interaction). Note that all independent variables were based on focal users' first three months of online activity.

(1) *#INT_w/focal* and *#INT_w/seed* are the total numbers of interactions a focal user had with other focal users and seed users, respectively.

(2) We further distinguished active interactions from support-receiving interactions and identified four independent variables: active interactions with other focal users (*#AI_focal),* support received from other focal users (*#SRI_focal),* active interactions with seed users (*#AI_seed*) and support received from seed users (*#SRI_seed*).

(3) To examine the effects of different types of seed users, we divided seed users into two subgroups according to their quit status at the start of study: current smokers and ex-smokers. Two additional variables-*#SRI_seedc* and *#SRI_seedx*-measure the amount of support focal users received from seed users who were current smokers and who were ex-smokers respectively.

Control variables included demographic and smoking characteristics collected in the baseline survey. These binary variables included whether the focal user was married (*MaritalStatus*) (Kim et al., 2017; Mermelstein et al., 1986; Spiegel et al., 1993), whether the focal user was living with smokers (*Live_w/Smokers*), *Gender* (Graham et al., 2017), whether the focal user was Hispanic or Latino (*Ethnicity*) (Graham et al., 2017), whether the focal user was employed (*WorkStatus*) (Graham et al., 2017), and the degree of nicotine dependence (*NicDependency*) (Kim et al., 2017).

Table 1 lists all variables in our model and their summary statistics. No Pearson correlation coefficients among variables were higher than 0.6 and all Variance Inflation Factor (VIF) values were lower than 2, suggesting low-to-moderate multicollinearity. Because the distributions of variables related to the number of OHC interactions were highly skewed, we used log-transformed values of these variables. Then, we standardized all control variables and independent variables with Z-scores, which helped us better observe how the changes on independent variables affected the outcome (i.e., focal users' 7-day abstinence status at 3-months post enrollment).

Table 1. Descriptive Statistics (N =188).

Variable	Description	Mean	Sd	(Min, Max)
Abstinence	Whether the focal user was abstinent for 7-days at the 3-month follow-up	0.18	0.39	(0, 1)
MaritalStatus	Whether the focal user was married	0.19	0.39	(0, 1)

Live_w/Smokers	Whether the focal user was living with smokers	0.29	0.45	(0, 1)
Gender	Female (0) or male (1)	0.61	0.49	(0, 1)
Ethnicity	Whether the focal user was Hispanic or Latino	0.18	0.39	(0, 1)
WorkStatus	Whether the focal user was employed	0.15	0.36	(0, 1)
NicDependency	Nicotine dependence levels at baseline (low, low/moderate, moderate, high)	2.45	0.91	(1, 4)
#INT_w/focal	The number of interactions with other focal users	7.41	15.09	(0, 103)
#INT_w/seed	The number of interactions with seed users	13.72	21.78	(0, 163)
#AI_focal	Active interactions with other focal users	3.55	10.84	(0, 78)
#SRI_focal	Social support received from other focal users	3.86	8.61	(0, 52)
#AI_seed	Active interactions with seed users	2.32	8.63	(0, 76)
#SRI_seed	Social support received from seed users	11.39	15.87	(0, 102)
#SRI_seedc	Social support received from seed users who were current smokers	0.19	0.86	(0, 9)
#SRI_seedx	Social support received from seed users who were ex-smokers	11.21	15.49	(0, 99)

Overall, our basic logistic regression model (Model 1) can be represented as follows:

$$ln\left(\frac{P(DV=1)}{1-P(DV=1)}\right) = \beta_0 + \beta_1 \cdot MaritalStatus + \beta_2 \cdot Live\_w/Smokers + \beta_3 \cdot Gender + \beta_4 \cdot Ethnicity + \beta_5 \cdot WorkStatus + \beta_6 \cdot NicDependency + \beta_7 \cdot \#INT\_w/focal + \beta_8 \cdot \#INT\_w/seed\_users + \varepsilon \quad \text{(Eq-1)}$$

where β_1, ... β_8 are the coefficients, β_0 is a constant, and ε is the error term.

Results

Table 2 shows results of four different models. *First*, the positive and significant effect of *#INT_w/seed* in Model 1 suggests that focal users who have more interactions with seed users are more likely to stop smoking. *Second*, Model 2 dissects interactions with seed and focal users and

finds that support received from seed users is the only significant and positive predictor for focal users' abstinence. *Last*, Model 3 examines the effect of seed users with different smoking statuses and finds that only support from seed users who were abstinent positively and significantly predicted focal user abstinence. In other words, focal users receiving a higher level of support from seed users who were former smokers were more likely to quit, compared to focal users who received less support from those seed users. All four models yielded consistent findings of control variables. For instance, *MaritalStatus* was a positive and significant predictor and *Ethnicity* was marginally significantly associated with cessation.

Table 2. Results of Logistic Regression Models.

	Model 0	Model 1	Model 2	Model 3
Intercept	-1.893** (0.697)	-1.790* (0.702)	-1.834** (0.706)	-1.850** (0.711)
Control variables				
MaritalStatus	1.208** (0.462)	1.413** (0.503)	1.339** (0.512)	1.387** (0.523)
Live_w/Smokers	-0.250 (0.470)	-0.367 (0.485)	-0.291 (0.499)	-0.286 (0.509)
Gender	-0.074 (0.402)	-0.183 (0.421)	-0.207 (0.425)	-0.154 (0.430)
Ethnicity	0.802. (0.455)	0.910. (0.466)	0.875. (0.468)	0.900. (0.471)
WorkStatus	0.587 (0.502)	0.450 (0.514)	0.434 (0.519)	0.405 (0.525)
NicDependency	-0.033 (0.223)	-0.074 (0.224)	-0.06 (0.224)	-0.084 (0.227)
Independent variables				
Focal users' interactions with whom				
log(#INT_w/focal)		-0.296 (0.235)		
log(#INT_w/seed)		0.438* (0.214)		
Received or provided social support				
log(#AI_focal)			-0.263 (0.243)	-0.244 (0.247)
log(#SRI_focal)			-0.023 (0.232)	0.032 (0.236)
log(#AI_seed)			-0.203 (0.258)	-0.118 (0.268)
log(#SRI_seed)			0.546* (0.246)	

log(#SRI_seedc)				-0.329 (0.290)
log(#SRI_seedex)				0.573* (0.246)
AIC	180.035	179.475	181.674	181.838
Note: (1) *** $p < .001$, ** $p < .01$, * $p < .05$ (2) SEs are in parentheses.				

To further examine how seed users with different smoking status (i.e., current vs former smokers) help focal users in the OHC, we compared linguistic characteristics between replies from seed users who were former smokers and replies from seed users who were current smokers. Linguistic Inquiry and Word Count (LIWC) has been widely used to capture linguistic features from text data. LIWC categorizes words into around 90 groups (Pennebaker et al., 2015). To assess language use differences, we conducted Mann-Whitney U tests on the two groups of seed users for each word category (i.e., ex-smokers vs. current smokers). Results are shown in Table 3. Results showed that seed users who were ex-smokers wrote longer replies with more complex structures (e.g., more conjunctions). Those longer replies provided space for ex-smoker seed users to be more assertive (e.g., more words related to certainty) and talk more about the rewards of cessation. In contrast, seed users who were current smokers used more third-person pronouns when supporting focal users.

Table 3. Linguistic Characteristics of Support from Seed Users with Different Smoking Status.

		Support from seed users who were ex-smokers		Support from seed users who were current smokers	
	p-value	Mean	SD	Mean	SD
Word Count	*	42.22	44.78	25.92	27.69
Pronoun: she/he	*	0.19	1.16	0.81	2.52
Conjunction	*	5.38	4.33	2.13	3.54
Negations	*	1.60	2.67	0.19	0.48
Certainty	*	1.93	2.99	0.51	1.85
Reward	*	3.48	4.81	1.21	2.10
Note: * *p < 0.05; ** p < 0.005; *** p < 0.001*					

Summary and future work

Our study has two main findings: *First*, social support from seed users is indeed a predictor for focal users' abstinence. While the effect of interactions with peers was not significant in the model, this does not necessarily suggest that peer support from other focal users in an OHC is not beneficial for better health outcomes. Instead, these results suggest that for a new OHC with a small user base, the role of seed users at this early stage can be very important not only in generating more online participation but also in promoting health behavior change (i.e., abstinence). *Second*, social support from seed users who have successfully quit better predicts focal users' abstinence than support from seed users who are still smoking. We conjecture that seed users who already quit may be better able to share successful strategies and better motivate others based on the rewards of abstinence. These findings can inform strategies to jumpstart an OHC by seeding the community with users who have successful experience in achieving the desirable health outcome. The language patterns that we discovered can be used to coach seed users how to better seed users.

There are several avenues for future research. For instance, more fine-grained analyses of online interaction content could further help us better understand the nature of support that prompts abstinence. Developing machine learning-based predictive models for abstinence based on online interactions would be interesting as well.

Reference

Baskerville, N. B., Azagba, S., Norman, C., McKeown, K., & Brown, K. S. (2015). Effect of a digital social media campaign on young adult smoking cessation. *Nicotine & Tobacco Research*, *18*(3), 351–360.

Bishop, J. (2007). Increasing participation in online communities: A framework for human–computer interaction. *Computers in Human Behavior*, *23*(4), 1881–1893.

Centers for Disease Control and Prevention. (n.d.). *Benefits of Quitting*. 2020.

Centers for Disease Control and Prevention. (2021). *Smoking & Tobacco Use*.

Cheung, Y. T. D., Chan, C. H. H., Lai, C.-K. J., Chan, W. F. V., Wang, M. P., Li, H. C. W., Chan, S. S. C., & Lam, T.-H. (2015). Using WhatsApp and Facebook online social groups for smoking relapse prevention for recent quitters: a pilot pragmatic cluster randomized controlled trial. *Journal of Medical Internet Research*, *17*(10), e4829.

Cobb, N. K., Graham, A. L., & Abrams, D. B. (2010). Social network structure of a large online community for smoking cessation. *American Journal of Public Health*, *100*(7), 1282–1289.

Cohen, S. (1988). Psychosocial models of the role of social support in the etiology of physical disease. *Health Psychology*, *7*(3), 269.

Graham, A. L., & Amato, M. S. (2019). Twelve million smokers look online for smoking cessation help annually: Health Information National Trends Survey data, 2005–2017. *Nicotine and Tobacco Research*, *21*(2), 249–252.

Graham, A. L., Zhao, K., Papandonatos, G. D., Erar, B., Wang, X., Amato, M. S., Cha, S., Cohn, A. M., & Pearson, J. L. (2017). A prospective examination of online social network dynamics and smoking cessation. *PloS One*, *12*(8), e0183655.

Iriberri, A., & Leroy, G. (2009). A life-cycle perspective on online community success. *ACM*

Computing Surveys (CSUR), *41*(2), 1–29.

J. Baxter, O. (2017). *Prescription drug abuse increasingly seen as a major U.S. public health problem*. Pew Research Center.

Kim, S. S., Darwish, S., Lee, S. A., & DeMarco, R. F. (2017). A pilot study of a smoking cessation intervention for women living with HIV: study protocol. *Open Access Journal of Clinical Trials*, *9*, 11–20.

Lichtenstein, E., Glasgow, R. E., & Abrams, D. B. (1986). Social support in smoking cessation: In search of effective interventions. *Behavior Therapy*, *17*(5), 607–619.

McLeish, J., & Redshaw, M. (2015). Peer support during pregnancy and early parenthood: a qualitative study of models and perceptions. *BMC Pregnancy and Childbirth*, *15*(1), 1–14.

Mdodo, R., Frazier, E. L., Dube, S. R., Mattson, C. L., Sutton, M. Y., Brooks, J. T., & Skarbinski, J. (2015). Cigarette smoking prevalence among adults with HIV compared with the general adult population in the United States: cross-sectional surveys. *Annals of Internal Medicine*, *162*(5), 335–344.

Mermelstein, R., Cohen, S., Lichtenstein, E., Baer, J. S., & Kamarck, T. (1986). Social support and smoking cessation and maintenance. *Journal of Consulting and Clinical Psychology*, *54*(4), 447.

Myneni, S., Fujimoto, K., Cobb, N., & Cohen, T. (2015). Content-driven analysis of an online community for smoking cessation: integration of qualitative techniques, automated text analysis, and affiliation networks. *American Journal of Public Health*, *105*(6), 1206–1212.

Pechmann, C., Delucchi, K., Lakon, C. M., & Prochaska, J. J. (2017). Randomised controlled trial evaluation of Tweet2Quit: a social network quit-smoking intervention. *Tobacco Control*, *26*(2), 188–194.

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*.

Ploderer, B., Smith, W., Howard, S., Pearce, J., & Borland, R. (2013). Patterns of support in an online community for smoking cessation. *Proceedings of the 6th International Conference on Communities and Technologies*, 26–35.

Resnick, P. J., Janney, A. W., Buis, L. R., & Richardson, C. R. (2010). Adding an online community to an internet-mediated walking program. Part 2: strategies for encouraging community participation. *Journal of Medical Internet Research*, *12*(4), e72.

Solomon, J., & Wash, R. (2012). Bootstrapping wikis: developing critical mass in a fledgling community by seeding content. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, 261–264.

Spiegel, D., Frischholz, E. J., Fleiss, J. L., & Spiegel, H. (1993). Predictors of smoking abstinence following a single-session restructuring intervention with self-hypnosis. *American Journal of Psychiatry*, *150*, 1090.

Stearns, M., Nambiar, S., Nikolaev, A., Semenov, A., & McIntosh, S. (2014). Towards evaluating and enhancing the reach of online health forums for smoking cessation. *Network Modeling Analysis in Health Informatics and Bioinformatics*, *3*(1), 1–11.

Wagener, T. L., Floyd, E. L., Stepanov, I., Driskill, L. M., Frank, S. G., Meier, E., Leavens, E. L., Tackett, A. P., Molina, N., & Queimado, L. (2017). Have combustible cigarettes met their match? The nicotine delivery profiles and harmful constituent exposures of second-generation and third-generation electronic cigarette users. *Tobacco Control*, *26*(e1), e23–e28.

Wang, Xi, Zhao, K., & Street, N. (2017). Analyzing and predicting user participations in online

health communities: a social support perspective. *Journal of Medical Internet Research*, *19*(4), e130.

Wang, Xiangyu, High, A., Wang, X., & Zhao, K. (2021). Predicting users' continued engagement in online health communities from the quantity and quality of received support. *Journal of the Association for Information Science and Technology*, *72*(6), 710–722.

Wright, K. B., Bell, S. B., Wright, K. B., & Bell, S. B. (2003). Health-related support groups on the Internet: Linking empirical findings to social support and computer-mediated communication theory. *Journal of Health Psychology*, *8*(1), 39–54.

Young, C. (2013). Community management that works: how to build and sustain a thriving online health community. *Journal of Medical Internet Research*, *15*(6), e119.

Zhang, M., & Yang, C. C. (2015). Using content and network analysis to understand the social support exchange patterns and user behaviors of an online smoking cessation intervention program. *Journal of the Association for Information Science and Technology*, *66*(3), 564–575.

Zhao, K., Wang, X., Cha, S., Cohn, A. M., Papandonatos, G. D., Amato, M. S., Pearson, J. L., & Graham, A. L. (2016). A multirelational social network analysis of an online health community for smoking cessation. *Journal of Medical Internet Research*, *18*(8), e5985.