

# Diffusion of an Evidence-Based Smoking Cessation Intervention Through Facebook: A Randomized Controlled Trial


Nathan K. Cobb, MD, Megan A. Jacobs, MPH, Paul Wileyto, PhD, Thomas Valente, PhD, and Amanda L. Graham, PhD

**Objectives.** To examine the diffusion of an evidence-based smoking cessation application (“app”) through Facebook social networks and identify specific intervention components that accelerate diffusion.

**Methods.** Between December 2012 and October 2013, we recruited adult US smokers (“seeds”) via Facebook advertising and randomized them to 1 of 12 app variants using a factorial design. App variants targeted components of diffusion: duration of use ( $t$ ), “contagiousness” ( $\beta$ ), and number of contacts ( $Z$ ). The primary outcome was the reproductive ratio ( $R$ ), defined as the number of individuals installing the app (“descendants”) divided by the number of a seed participant’s Facebook friends.

**Results.** We randomized 9042 smokers. App utilization metrics demonstrated between-variant differences in expected directions. The highest level of diffusion ( $R=0.087$ ) occurred when we combined active contagion strategies with strategies to increase duration of use (incidence rate ratio=9.99; 95% confidence interval=5.58, 17.91;  $P<.001$ ). Involving nonsmokers did not affect diffusion.

**Conclusions.** The maximal  $R$  value (0.087) is sufficient to increase the numbers of individuals receiving treatment if applied on a large scale. Online interventions can be designed a priori to spread through social networks. (*Am J Public Health.* 2016;106:1130–1135. doi:10.2105/AJPH.2016.303106)

 See also Galea and Vaughan, p. 973.

Given tobacco’s deadly and addictive properties, there continues to be a pressing need for effective and broadly disseminated tobacco cessation interventions. Most smokers attempt to quit without any form of assistance,<sup>1,2</sup> and tobacco use rates have remained relatively static over the past 5 years at around 20%.<sup>3</sup> There have been numerous efforts to build effective online interventions for tobacco cessation.<sup>4,5</sup> Internet smoking cessation programs have the potential to reach a large number of people in a cost-efficient manner, but their uptake is often tied to expensive marketing and promotion efforts.<sup>6</sup>

The literature on diffusion of innovations suggests an alternative model for treatment dissemination in which an intervention proliferates through an existing social network, “virally” spreading from smoker to smoker.<sup>7</sup> Many studies have documented the role of interpersonal influence in behavior

change,<sup>8–12</sup> including the viral spread of cessation in large networks.<sup>13</sup> The explosion of online social networks, such as Facebook, has potentially changed the concept of how network-based interventions might function by shifting interpersonal influence to mediated communications platforms.<sup>14</sup> The broad diffusion of Facebook applications (“apps”) and games has become part of the national

discourse, making the phrase “to go viral” part of the vernacular. Viral applications are generally disseminated either by active mechanisms (e.g., via invitations, e-mail) or passive mechanisms (e.g., via observation).

As with the spread of an infectious disease, the diffusion of online social network applications can be modeled with the reproductive ratio ( $R$ ), which is determined by the duration that an individual remains “infectious” ( $t$ ), the relative “contagiousness” of the application itself ( $\beta$ ), and the number of contacts ( $Z$ ).

$$(1) R = t\beta Z$$

In cases in which the mean value of  $R$  exceeds 1.0 (i.e., when a typical user spreads the application to more than 1 person), an application will diffuse exponentially through a network. When  $R$  is less than but close to 1.0, diffusion depends on recurring new “infections” but can spread across multiple generations. For example, an application with an  $R$  value of 0.9 will reach an average of 6 additional participants when traced out to 10 generations. For online applications with rapid diffusion cycles (days or even hours),  $R$  values well below 1.0 still have the potential to promulgate an intervention. The concept of the reproductive rate is common across network science, public health, epidemiology, and social application design, making it

## ABOUT THE AUTHORS

Nathan K. Cobb is with the Department of Pulmonary and Critical Care, Georgetown University Medical Center, Washington, DC, and the Department of Health, Behavior and Society, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD. Megan A. Jacobs is with the Schroeder Institute for Tobacco Research and Policy Studies, Truth Initiative, Washington, DC. Paul Wileyto is with the Department of Biostatistics & Epidemiology, University of Pennsylvania School of Medicine, Philadelphia. Thomas Valente is with the Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles. Amanda L. Graham is with the Schroeder Institute for Tobacco Research and Policy Studies, Truth Initiative, Washington, DC, and the Department of Oncology, Georgetown University Medical Center/Cancer Prevention and Control Program, Lombardi Comprehensive Cancer Center, Washington, DC.

Correspondence should be sent to Amanda L. Graham, PhD, Schroeder Institute for Tobacco Research and Policy Studies, Truth Initiative, 900 G St, NW, 4th Fl, Washington, DC 20001 (e-mail: agraham@truthinitiative.org). Reprints can be ordered at <http://www.ajph.org> by clicking the “Reprints” link.

This article was accepted January 29, 2016.

doi: 10.2105/AJPH.2016.303106

a unifying metric that can be used both for intervention design and evaluation while also encouraging cross-program comparisons.

We hypothesized that the Facebook social network could be used to virally disseminate an evidence-based smoking cessation application. We constructed a Facebook app that varied specific features of the 3 components of  $R$ : (1) the length of time ( $t$ ) participants used the application, (2) how aggressive the application was in promoting itself to a participant's friends ( $\beta$ , or contagiousness), and (3) the number of Facebook friends ( $Z$ , or contacts). The app was designed so that these components of  $R$  (time, contagiousness, contacts) could be enabled or disabled for individual participants and evaluated in an experimental design. The rate of intervention diffusion ( $R$ ) was the outcome of interest.

## METHODS

The trial employed a fractional factorial design to evaluate the impact of  $t$  (time),  $\beta$  (contagiousness), and  $Z$  (contacts). Because contagiousness is particularly noteworthy, we implemented both a passive and active version of  $\beta$ . These 4 components of the app were varied so that features were turned on or off or set at high or low, resulting in 16 cells or conditions. Four of these 16 cells, however, had no theoretical potential for diffusion (e.g., the cells with time or contagion turned off) and we excluded them from the trial, resulting in a 12-cell fractional factorial design. Analyses accommodated the fact that the fractional factorial design included the interaction cell (i.e., all features turned on) but not the true origin cell (i.e., all features turned off). We identified all features in a theory-driven ideation and testing phase designed to map each one back to the identified variables in an effort to maximize the likelihood of diffusion. This mapping was somewhat imperfect, reflecting a real-world design process that by necessity considered the feasibility of implementing each one within the Facebook app environment and its familiarity to participants.

We varied the duration of exposure ( $t$ ) through 2 features designed to increase the amount of time that participants used the application: expanded content and proactive messaging. We provided participants with

topic-specific quit guides, videos, and animated content beyond the core intervention, as well as Facebook reminders encouraging return visits. We implemented contagiousness ( $\beta$ ) in 2 forms:  $\beta$ -active, where contact with friends was proactively initiated by the user through the app, and  $\beta$ -passive, where friends were exposed to information about the participant's behavior automatically. Participants with  $\beta$ -active could directly invite friends to install the application and share content so that it appeared in the Facebook News Feed of their friends. Participants were also exposed to normative data comparing them with other users on use metrics and cessation milestones. For participants with  $\beta$ -passive turned on, the application automatically posted content to their News Feed where friends might see it. All posts generated opportunities for friends to like, comment on, share the application-generated object, or click on the shared content. We varied the number of contacts ( $Z$ ) by creating a version of the application to allow nonsmokers to provide support for smoking cessation. Participants randomized to cells with  $Z$  enabled were encouraged to have nonsmoking friends install the application to support their cessation efforts. Although the levels for each feature were binary (i.e., features were turned on or off), the resulting measurements and variables were all continuous. For example, a participant may have been randomized to be able to invite friends or not (binary), but the number of friends they actually invited was recorded and analyzed as a continuous variable.

## Recruitment and Enrollment

The trial was conducted entirely within Facebook between December 2012 and October 2013 (ClinicalTrials.gov identifier: NCT01746472). All participants were registered users of Facebook. We recruited initial adopters ("study seeds") via Facebook advertising and earned media. Individuals that clicked through to the application were shown a Facebook dialog box asking for installation permission followed by informed consent. Inclusion criteria for study seeds were US residency, current smoking, age 18 years or older, an active English-language Facebook account and e-mail address, and acceptance of Facebook permissions for application installation. The only exclusion

criterion was having a Facebook friend that had already installed the application. We determined age, number of existing friends that were already application users, and location-related eligibility via Facebook metrics at installation; smoking status was determined via self-report immediately after informed consent. Users excluded from study participation for any reason were allowed access to the Facebook application.

To track diffusion paths, we embedded tags within all links to the application and in content shared by users. We identified new users that reached the application through an existing study seed in real time and excluded them from becoming seeds themselves. "Descendants" were users—both smokers and nonsmokers—that could be tied to an existing participant through a tracking mechanism. Descendants installed the application and accepted informed consent in the same manner as seed users.

To further characterize the sample beyond the limited metrics available via Facebook, we subsampled 10% of seed participants to complete a brief Web-based survey at baseline and 30 days after enrollment. Survey participants were reimbursed US \$20 for each completed survey. Users randomized to the subsample who did not complete the baseline survey were able to use the application but were excluded from the study.

## Randomization and Intervention

We randomized seed users to 1 of the 12 cells using an adaptive "biased-coin" strategy.<sup>15</sup> We assigned descendants to the same cell as the "parent" seed user; when the parent was unclear, we assigned descendants to the same cell as the friend that installed the application most recently.

Details about the development of the app and the trial design have been described in detail previously.<sup>16</sup> Briefly, the Facebook intervention is based on the US Public Health Service (PHS) Guideline "5As" model (Ask, Advise, Assess, Assist, and Arrange).<sup>17</sup> A cartoon physician named Dr. Youkwitz asks participants if they smoke and advises smokers to quit. He then assesses their readiness to make a quit attempt and assists them by providing a tool ("Quit Date Wizard") for planning a quit attempt and setting a quit date. If a user sets a quit date, the application

displays a countdown to that date or an estimate of savings since that date (money saved, estimate of life saved). Users who do not set a quit date in their first visit may set one at any time. The application arranges follow-up in the form of daily check-ins with Dr. Youkwitz that provide tailored, personalized information and support and gather smoking status. Users randomized to cells that have the components of *t* enabled receive proactive Facebook application requests alerting them that a check-in is ready for them in the application. Participants are prompted at each check-in to confirm their quit date or update their smoking status. Smokers who have not set a quit date receive a variety of daily check-ins that include prompts to set a quit date, as well as evidence-based content incorporating the “5 Rs” (Relevance, Risk, Rewards, Roadblocks, and Repetition) derived from the PHS Guideline.

## Data Collection and Measures

Data collection occurred primarily through Facebook’s application programming interface (API). The API allowed our systems to interact directly with Facebook’s database to retrieve data about individual users (e.g., age, gender) and their immediate social network upon study enrollment. A Facebook user could learn about the application through paid advertising or earned media, word of mouth or observation of others’ use of the application within Facebook, or direct invitation. We used standardized mechanisms within the Facebook API to record “invitations,” Facebook timeline or wall posts, and subsequent “acceptance” or click-throughs by individuals, allowing a precise calculation of the number of descendants, or *R*. Among participants who could invite nonsmoking friends as supporters, we tracked out diffusion an additional degree to calculate the number of smokers recruited using nonsmokers as a bridge. We calculated *R* directly as the number of new smokers recruited by a single individual.

Using NetworkX,<sup>18</sup> we calculated ego-level social network metrics hypothesized to affect diffusion, including degree (the number of friends as reported by Facebook), the average local clustering coefficient (how tightly knit the personal network), and

betweenness centrality (the relative importance of individuals within their own network, normalized for degree). The system automatically recorded page views and interactions between the user and the application into a relational database. Utilization metrics derived from these events included return visits, specific pages viewed, and use of application tools. Additionally, we recorded social interactions via Facebook, including invitations, News Feed posts, and wall posts as well as interactions with system notifications and requests.

## Power Analysis

We determined sample size using an estimate of *R* based on metrics from viral online marketing campaigns, where *R* has ranged from 0.041 to 2.<sup>19,20</sup> Given the difficulty of estimating the required sample size for a primary outcome of diffusion, we followed a common-sense approach proposed by Watts and Dodds,<sup>20</sup> which involved seeding the application to as many individuals as possible rather than a carefully targeted few. The current study was powered for an *R* value below 1 at the first generation from seed users to guarantee productive analysis even if the intervention did not reach “viral threshold.” A sample size of 8000 seed participants provided 88% power to examine all between-factor analyses with a minimal detectable difference in basic reproductive rate of 0.1.

## Statistical Analyses

We conducted analyses using Stata version 13.0 (StataCorp LP, College Station, TX). We analyzed the primary outcome (*R*) with the generalized linear model (GLM) using a negative binomial error family to examine the counts of new cases that arose directly from each primary seed participant. The regression contained binary categorical predictors for *t* (time: 0, 1) and *Z* (expanded contact: 0, 1) and a 3-level categorical predictor for contagion mode that was  $\beta$ -active,  $\beta$ -passive, or  $\beta$ -active +  $\beta$ -passive (0, 1, 2). The model constant provided a measure of *R* for the reference group. Additional regression parameters are expressed as rate ratios, which multiply the baseline *R* value from the reference group. We tested a priori hypotheses against 0 using the log of the rate ratio

compared to the Wald  $\chi^2$  or *z*-score. We examined additional network metrics as covariates, including degree, clustering coefficient, count of the number of Facebook friends that knew another friend within the participant’s friend list, the sum of the number of friends shared between each friend and the participant, and multiple measures of centrality and transitivity. We screened variables for entry to the model and retained them if  $P \leq .2$ . We then added the covariates into the model using a forward stepwise procedure, with *P* values set at 0.1 to enter and 0.2 to remove.

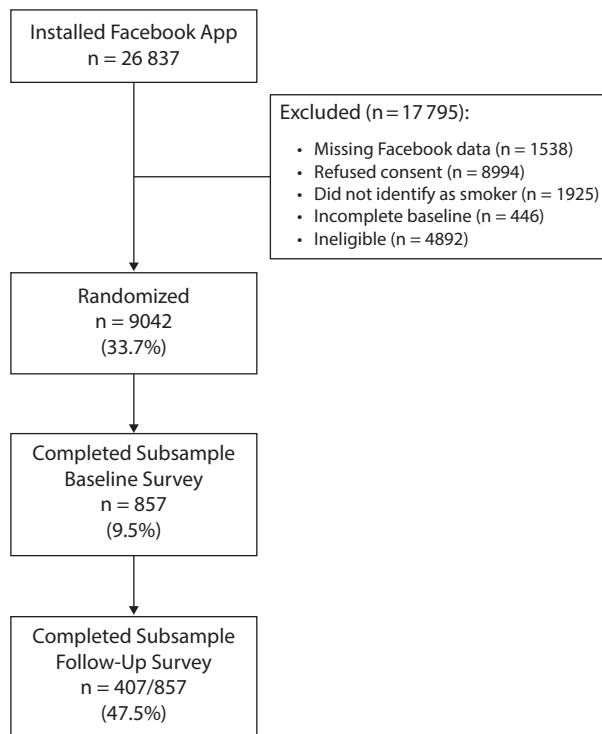
## RESULTS

Facebook recruitment ads were seen by approximately 54 million users (per Facebook’s advertising system): 26 837 individuals installed the application, of which 9042 met eligibility criteria, provided informed consent, and were randomized as seed users. The study CONSORT diagram is shown in Figure 1.

The 9042 seed users were mostly female (70%) with a mean age of 43.9 years (SD = 14.1). On average, participants had 307.8 (SD = 401.8) Facebook friends. There were no significant differences between cells for available demographic variables or network metrics. Among the subsample of survey participants (*n* = 857), 11% were non-White, 4% were Hispanic, 67% were married, and 58% had at least some college education.

## Application Use

Application use data are presented in Table 1. Participants randomized to cells with *t* (time) features enabled returned to the site more often (application visits:  $3.21 \pm 6.15$  vs  $1.91 \pm 3.62$ ;  $P < .001$ ) and saw more content (page views:  $29.02 \pm 37.96$  vs  $21.62 \pm 23.39$ ;  $P < .001$ ) than participants in cells with *t* disabled. Individuals with *Z* (expanded contacts) enabled were encouraged to reach out to nonsmoking friends; the number of nonsmoking descendants was slightly higher among participants with *Z* enabled than among those with *Z* disabled ( $0.022 \pm 0.17$  vs  $0.016 \pm 0.14$ ;  $P = .047$ ). Individuals in  $\beta$ -active cells invited an average of 3.51



Note. App = application.

FIGURE 1—Study CONSORT Diagram

(SD = 10.20) friends and posted application-related content to the walls of 0.63 (SD = 1.65) friends;  $\beta$ -passive participants made an average of 0.62 (SD = 2.34) posts to their own News Feed and 1.42 (SD = 3.04) to their own timeline.

TABLE 1—Mean Number of Times US Smokers Used Different Features (Utilization Metrics) by Factor From the Experimental Design (On or Off, High or Low): Facebook, December 2012–October 2013

Factor and Utilization Metric	On: High Condition, Mean (SD)	Off: Low Condition, Mean (SD)	<i>P</i>
<i>t</i>			
Application visits	3.21 (6.15)	1.91 (3.62)	<.001
Page views	29.02 (37.96)	21.62 (23.39)	<.001
Distinct pages viewed	14.40 (9.91)	12.05 (4.76)	<.001
Crave button “pushes”	2.47 (7.95)	1.88 (6.50)	<.001
Quit guides read	0.62 (1.60)	0.21 (0.40)	<.001
Quit dates set	0.89 (0.75)	0.84 (0.64)	<.001
Z: nonsmoking descendants	0.022 (0.17)	0.016 (0.14)	.047
$\beta$ -active			
Friends invited	3.51 (10.20)	NA	NA
Posts to friends’ feeds	0.79 (2.79)	NA	NA
Friends receiving feed posts	0.63 (1.65)	NA	NA
$\beta$ -passive			
Posts to own News Feed	0.62 (2.34)	NA	NA
Posts to own timeline	1.42 (3.04)	NA	NA

Note. *t* (time): expanded content, proactive contact; Z: nonsmoker supporters;  $\beta$ -active: invites, social comparison;  $\beta$ -passive: passive content sharing. NA = not applicable.

### Diffusion Outcome

The negative binomial model provides a baseline estimate of *R* from the model constant ( $R = 0.016$ ; 95% confidence interval [CI] = 0.011, 0.024;  $P < .001$ ). This is the *R* value calculated for the (imaginary) reference group (Cell 000), which has no enhanced content, does not allow nonsmokers as supporters, and has only passive contagion ( $\beta$ -passive). Marginal estimates of *R* values by cell are shown in Table 2. The effects of passive and active contagion are, by the nature of the fractional design, considered additive.

As shown in Table 3, there was an interaction between *t* (time) and  $\beta$  (contagion type; Wald  $\chi^2_2 = 6.73$ ;  $P = .035$ ). Compared with the reference group, *t* combined with  $\beta$ -passive reduced *R* (incidence rate ratio [IRR] = 0.475; 95% CI = 0.238, 0.949;  $P = .035$ ). There were main effects for  $\beta$ -active (IRR = 3.721; 95% CI = 2.379, 5.818;  $P < .001$ ) and  $\beta$ -active +  $\beta$ -passive (IRR = 4.047; 95% CI = 2.598, 6.305;  $P < .001$ ), both of which increased diffusion. Time (*t*) increased *R* in the presence of  $\beta$ -active by a factor of 2.486 (95% CI = 1.176, 5.255;  $P = .017$ ) and increased *R* in the presence of  $\beta$ -active +  $\beta$ -passive by a factor of 2.652 (95% CI = 1.262, 5.574;  $P = .010$ ). There was no significant effect of allowing nonsmoker supporters (*Z*) to use the application (IRR = 1.051; 95% CI = 0.872, 1.268;  $P = .585$ ). The model was very similar after adjustment for covariates.

The coefficients for the main effects of contagion type ( $\beta$ -active and  $\beta$ -active +  $\beta$ -passive) were similar ( $P = .57$ ). Likewise, the interaction terms were not significantly different from each other ( $P = .75$ ). The implication is that adding passive contagion to active contagion has no effect. For that reason, it is possible to collapse the contagion factor into a binary variable that represents “has active contagion” versus “has passive only,” randomized 2:1. For the unadjusted model, we have a reference group of passive-only contagion ( $\beta$ -passive, low *t*, and low *Z*), with an *R* value of 0.016 (95% CI = 0.011, 0.024;  $P < .001$ ). The interaction is similar to that in the model with 3-level contagion described in the previous paragraph. Adding enhanced content (*t* high) to passive-only contagion reduces the diffusion rate (IRR = 0.475; 95%

**TABLE 2—Estimates of R From the Unadjusted Model and Mean Number of Descendants From Data by Factorial Cell: Facebook, December 2012–October 2013**

Cell No.	Cell Size	t	Z	β-Active	β-Passive	R Model (95% CI)	R: Means From Data (95% CI)
1	0	0	0	0	0	NA	NA
2	749	0	0	0	1	0.016 (0.010, 0.023)	0.020 (0.008, 0.033)
3	750	0	0	1	0	0.060 (0.047, 0.074)	0.069 (0.046, 0.092)
4	761	0	0	1	1	0.066 (0.051, 0.080)	0.059 (0.041, 0.077)
5	0	0	1	0	0	NA	NA
6	752	0	1	0	1	0.017 (0.010, 0.024)	0.013 (0.001, 0.026)
7	751	0	1	1	0	0.064 (0.049, 0.078)	0.055 (0.038, 0.072)
8	752	0	1	1	1	0.069 (0.054, 0.084)	0.076 (0.055, 0.096)
9	0	1	0	0	0	NA	NA
10	763	1	0	0	1	0.008 (0.003, 0.012)	0.008 (0.001, 0.015)
11	752	1	0	1	0	0.071 (0.056, 0.087)	0.065 (0.046, 0.085)
12	751	1	0	1	1	0.083 (0.066, 0.100)	0.083 (0.050, 0.115)
13	0	1	1	0	0	NA	NA
14	754	1	1	0	1	0.008 (0.003, 0.013)	0.008 (0.002, 0.014)
15	751	1	1	1	0	0.075 (0.059, 0.091)	0.081 (0.056, 0.106)
16	756	1	1	1	1	0.087 (0.069, 0.105)	0.087 (0.063, 0.112)

Note. CI = confidence interval; NA = not applicable. t (time): expanded content, proactive contact; Z: nonsmoker supporters; β-active: invites, social comparison; β-passive: passive content sharing. Cells 1, 5, 9, and 13 were suppressed in the fractional factorial design.

CI = 0.238, 0.949; *P* = .035). Active contagion increased *R* when *t* was low (low *t*: IRR = 3.885; 95% CI = 2.550, 5.918; *P* < .001), and this effect was magnified when *t* was high (high *t*: IRR = 9.993; 95% CI = 5.577, 17.905; *P* < .001). There was

no significant effect of *Z* (nonsmokers as supporters) on initial diffusion at the first generation.

## DISCUSSION

Decades of work has demonstrated that behavior can spread through networks; however, understanding and potentially augmenting diffusion can be enhanced through online networks such as Facebook.<sup>21,22</sup> In this study, we manipulated individual features of a Facebook app to affect 3 hypothesized drivers of diffusion: *t* (time), *Z* (contacts), and β (contagion, both passive and active). The version of the app that enabled all features demonstrated an *R* value of 0.087, sufficient to augment diffusion and increase the numbers of individuals receiving treatment if applied at public health scale. Given the scale of social media, this level of diffusion indicates that the intervention may be cost beneficial.<sup>14</sup> Assuming that roughly 5% of the 42 million smokers in the United States<sup>3</sup> are willing to make a quit attempt and engage with a behavioral cessation intervention,<sup>1,2</sup> an *R* value of this magnitude may have a large impact on diffusion. For example, a public health campaign using a program similar to the one evaluated in this study designed to engage roughly 2.1 million smokers could reach an additional 182 700 participants (0.087 × 2.1 million) through viral spread, at no additional cost.

To our knowledge, this is the first study to show that intervention diffusion is a function of specific application elements. Studies of contagion typically infer interpersonal influence from correlational data.<sup>13</sup> Conversely, we manipulated intervention conditions and demonstrated that these variations had resulting effects on behavior within the application, including the likelihood that users spread it to others. Specifically, we found an interaction between the amount of content available (*t*, time) and whether participants could actively invite friends to the application (β-active, active contagion): users exposed to more content over a longer period of time actively spread the application to more contacts. By contrast, increasing the amount of content and contacts from the application (*t*) in conjunction with passive forms of diffusion (β-passive, passive contagion)

**TABLE 3—Incidence Rate (for the Reference Group) and Incidence Rate Ratios From Adjusted and Unadjusted Negative Binomial Models: Facebook, December 2012–October 2013**

Predictor	Unadjusted IRR <sup>a</sup> (95% CI)	Adjusted IRR (95% CI)
Reference group	0.016 (0.011, 0.024)	0.015 (0.010, 0.023)
t: expanded content	0.475 (0.238, 0.949)	0.468 (0.234, 0.935)
β-active	3.721 (2.379, 5.819)	3.768 (2.405, 5.902)
β-active+β-passive	4.047 (2.598, 6.305)	4.126 (2.644, 6.436)
Z: number contacts	1.051 (0.872, 1.268)	1.054 (0.873, 1.272)
β-active × t	2.486 (1.176, 5.255)	2.439 (1.152, 5.163)
β-active + β-passive × t	2.652 (1.262, 5.574)	2.619 (1.244, 5.513)
Covariates		
Age		0.792 (0.717, 0.876)
Count friends (degree)		0.819 (0.641, 1.047)
Total mutual friend count		0.917 (0.807, 1.041)
Count friends with friends		1.493 (1.208, 1.844)
Average clustering coefficient		1.118 (0.976, 1.281)
Ego betweenness centrality		1.142 (1.029, 1.267)
Graph transitivity		0.862 (0.760, 0.978)

Note. CI = confidence interval; IRR = incidence rate ratio. t (time): expanded content, proactive contact; Z: nonsmoker supporters; β-active: invites, social comparison; β-passive: passive content sharing.

<sup>a</sup>Negative binomial regression models.

appeared to decrease diffusion. It may be that by increasing the amount of shareable content we inadvertently decreased the perceived importance of each individual piece of information. Fewer pieces of content and more repeated sharing may be a better strategy to promote diffusion.

Encouraging participants to invite non-smoking friends did not affect diffusion. We hypothesized that nonsmokers might link clusters of smokers and facilitate viral spread. It is possible that participants did not invite nonsmokers or that nonsmoking friends failed to respond, perhaps perceiving that the app was not relevant to them. Future research should explore ways to involve nonsmokers before concluding that diffusion across clusters of smokers is not feasible.

There are 3 limitations of this study related to the design. First, the intervention was based on clinical guidelines for tobacco-dependence treatment. Incorporating individual treatment strategies into a network approach may have affected its capacity for diffusion. A second limitation is related to our implementation of the factorial design. We disabled cells that theoretically had no potential for diffusion, thereby decreasing the required sample size and recruitment costs. However, other mechanisms of contagion may have existed; it is possible that  $R$  in the 4 disabled cells would not have been zero as we presumed. Additionally, Facebook may have changed the algorithms that determined how information was presented to users in ways that we could not detect. These algorithms could have prevented users from seeing shared content at all, thus affecting the effectiveness of our  $\beta$ -passive strategies and diffusion seen from this factor. Finally, the use of  $R$  as the primary outcome is uncommon in the existing tobacco control literature. However, we deliberately chose it as a unifying metric appropriate for the entire life cycle of a program that prioritized dissemination from initial feature selection through evaluation. Other metrics may be more appropriate for different programs or studies; however,  $R$  provides a unique capacity to succinctly communicate the effect of an intervention to multiple stakeholders and audiences and provides a useful analogy for propagating and studying viral dissemination.

Our findings confirm that viral spread is indeed possible to manipulate within an

online application for health behavior change. Although our results may not be directly generalizable to other programs or software applications, we believe the general approach is innovative and merits further research. Using these findings to drive the population-level impact of behavioral interventions will require that intervention designers consider the importance of diffusion alongside metrics of efficacy. Interventions that are effective in modifying behavioral risk factors and that spread efficiently through social networks have the potential to affect millions of online network users. **AJPH**

### CONTRIBUTORS

N. K. Cobb, M. A. Jacobs, P. Wileyto, and A. L. Graham analyzed and interpreted data; they had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. N. K. Cobb, M. A. Jacobs, and A. L. Graham acquired data and supervised the study. N. K. Cobb and A. L. Graham conceptualized and designed the study and obtained funding. N. K. Cobb, P. Wileyto, and A. L. Graham provided statistical analysis. M. A. Jacobs and P. Wileyto provided administrative, technical, and material support. All authors contributed to the drafting of the article.

### ACKNOWLEDGMENTS

This research was funded by grant R01CA155369 from the National Cancer Institute of the National Institutes of Health.

**Note.** The funding agency had no role in the study design, analysis, interpretation of data, decision regarding publication, or preparation of this article.

### HUMAN PARTICIPANT PROTECTION

The study protocol was approved by Schulman Associates institutional review board (formerly Independent IRB).

### REFERENCES

- Shiffman S, Brockwell SE, Pillitteri JL, Gitchell JG. Use of smoking-cessation treatments in the United States. *Am J Prev Med.* 2008;34(2):102–111.
- Cokkinides VE, Ward E, Jemal A, Thun MJ. Underuse of smoking-cessation treatments: results from the national health interview survey, 2000. *Am J Prev Med.* 2005;28(1):119–122.
- Agaku IT, King BA, Husten CG, et al. Tobacco product use among adults—United States, 2012–2013. *MMWR Morb Mortal Wkly Rep.* 2014;63(25):542–547.
- Myung SK, McDonnell DD, Kazinets G, Seo HG, Moskowitz JM. Effects of web- and computer-based smoking cessation programs: meta-analysis of randomized controlled trials. *Arch Intern Med.* 2009;169(10):929–937.
- Shahab L, McEwen A. Online support for smoking cessation: a systematic review of the literature. *Addiction.* 2009;104(11):1792–1804.
- McCausland KL, Curry LE, Mushro A, Carothers S, Xiao H, Vallone DM. Promoting a web-based smoking cessation intervention: implications for practice. *Cases Public Health Commun Mark.* 2011;5:3–26.
- Valente TW. *Network Models of the Diffusion of Innovations.* Cresskill, NJ: Hampton Press; 1995.

- Iyengar R, Van den Bulte C, Valente TW. Further reflections on studying social influence in new product diffusion. *Mark Sci.* 2011;30(2):230–232.
- Myneni S, Iyengar S, Cobb NK, Cohen T. Identifying persuasive qualities of decentralized peer-to-peer online social networks in public health. In: Berkovsky S, Freyne J, eds. *Persuasive Technology.* Berlin, Germany: Springer; 2013:155–160.
- Cobb NK, Mays D, Graham AL. Sentiment analysis to determine the impact of online messages on smokers' choices to use varenicline. *J Natl Cancer Inst Monogr.* 2013;2013(47):224–230.
- Cobb NK, Graham AL, Abrams DB. Social network structure of a large online community for smoking cessation. *Am J Public Health.* 2010;100(7):1282–1289.
- Valente TW. *Social Networks and Health: Models, Methods, and Applications.* New York, NY: Oxford University Press; 2010.
- Christakis NA, Fowler JH. The collective dynamics of smoking in a large social network. *N Engl J Med.* 2008;358(21):2249–2258.
- Valente TW. Network interventions. *Science.* 2012;337(6090):49–53.
- Schulz KF, Grimes DA. Generation of allocation sequences in randomised trials: chance, not choice. *Lancet.* 2002;359(9305):515–519.
- Cobb NK, Jacobs MA, Saul J, Wileyto EP, Graham AL. Diffusion of an evidence-based smoking cessation intervention through Facebook: a randomised controlled trial study protocol. *BMJ Open.* 2014;4(1):e004089.
- Fiore MC, Jaén CR, Baker TB, et al. *Treating Tobacco Use and Dependence: 2008 Update.* Rockville, MD: US Dept of Health and Human Services, Public Health Service; May 2008.
- Hagberg AA, Schult DA, Swart PJ. Exploring network structure, dynamics, and function using NetworkX. Paper presented at: 7th Python in Science Conference; August 1, 2008; Pasadena, CA. Available at: <http://math.lanl.gov/~hagberg/Papers/hagberg-2008-exploring.pdf>. Accessed March 26, 2015.
- Kalyanam K, McIntyre S, Masonis JT. Adaptive experimentation in interactive marketing: the case of viral marketing at Plaxo. *J Interact Market.* 2007;21(3):72–85.
- Watts DJ, Dodds PS. Influentials, networks, and public opinion formation. *J Consum Res.* 2007;34(4):441–458.
- Aral S, Walker D. Identifying influential and susceptible members of social networks. *Science.* 2012;337(6092):337–341.
- Bond RM, Fariss CJ, Jones JJ, et al. A 61-million-person experiment in social influence and political mobilization. *Nature.* 2012;489(7415):295–298.