IMPLEMENTATION, POLICY AND COMMUNITY ENGAGEMENT
REVIEW ARTICLE

A review of social media methods and lessons learned from the National Children's Study

Amelia Burke-Garcia[1]*, Kate Winseck[2], Leslie Cooke Jouvenal[3], David Hubble[1]
and Kathryn M. Kulbicki[1]

[1] Westat, Rockville, MD, USA
[2] National Institutes of Health, Bethesda, MD, USA
[3] Formerly of the National Institutes of Health, Bethesda, MD, USA

* Address for correspondence: A. Burke-Garcia, M.A., Ph.D., Westat, 1600 Research Blvd, Rockville, MD 20850, USA.
(Email: ameliaburke-garcia@westat.com)

Introduction. Given the reach and influence of social media, the National Children's Study Vanguard Study evaluated the feasibility, acceptability, and cost of using social media to support participant retention.

Methods. We describe a social media experiment designed to assess the impact of social media on participant retention, discuss several key considerations for integrating social media into longitudinal research, and review factors that may influence engagement in research-related social media.

Results. User participation varied but was most active when at launch. During the short life of the private online community, a total of 39 participants joined. General enthusiasm about the prospect of the online community was indicated. There were many lessons learned throughout the process in areas such as privacy, security, and Institutional Review Board clearance. These are described in detail.

Conclusions. The opportunity to engage participants in longitudinal research using online social networks is enticing; however, more research is needed to consider the feasibility of their use in an ongoing manner. Recommendations are presented for future research seeking to use social media to improve retention in longitudinal research.

Introduction

Led out of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) at the National Institutes of Health (NIH), the National Children's Study (NCS) Vanguard Study was a pilot study, and precursor to the planned Main Study, which enrolled women and their children born in 40 study areas that represent the diverse geographic areas of the United States. The purpose of the Vanguard Study was to evaluate the feasibility, acceptability, and cost of recruitment and operations for the NCS Main Study. As planned, the NCS would have examined the impacts and interface of genetics and environment on the growth, development, and health of children across the United States, following them from before birth until age 21 years. Broadly defined, "environment" included the physical, chemical, biological, and psychosocial-cultural environments. Data were to be collected from pre-pregnancy until the children reached adulthood. Ultimately, it was decided that moving forward with the NCS Main Study was not feasible as planned and it was never launched. The Vanguard Study was closed in December 2014. Nonetheless, data from the pilot study can be used to address a number of research questions.

One such research question asked whether the use of social media could help support long-term participant retention in the NCS. Data suggest both broad and deep reach and influence of social media among the NCS Vanguard Study's target audience. For instance, 70% of Americans now use social media to connect with one another, engage with news content, and share information [1] and more than 80% of online Americans (and ~60% of all American adults) use the web to search for health information [2]. This increasing popularity of social media is having an important impact on health research. The use of different social media platforms (including Facebook and other interactive online programs) has been found to decrease attrition in longitudinal studies, encourage participants to engage in follow-up protocols, and generate active dialog in communities nationwide [3, 4]. In addition, social media can be used to reach very specific populations as its use transcends demographics [5]. However, while findings from this research suggest that social media offers new avenues for reaching and engaging research participants, implementation of these strategies can be fraught with challenges, key among which is security and confidentially of participant's personal information. Given this, the pilot study decided to evaluate social media as part of a larger outreach and engagement effort.

From September 2013 to the Study's close in December 2014, the NCS implemented various social media activities to engage the geographic communities where data collection was occurring. In this review article, we discuss the methods used to develop and implement a private, closed social media-like platform that was developed for the NCS. Finally, we present several implications and recommendations for future research seeking to use social media to improve participant retention in longitudinal research.

Literature Review

Social Media Use by Parents

Social media use has vastly increased in the last 10 years, with nearly two-thirds of American adults (65%) now using social networking sites [6]. Much of this use is focused on health-related information dissemination and engagement [6], as social media contribute to "facilitating, sharing, and obtaining health messages" [7]. In terms of health information, 86% of women report that they make the decisions about healthcare treatments for their entire family [8]; and research posits that health communicators should go where women spend time, which increasingly means online social network and blogging sites [9].

One recent study reports that 75% of parents use social networking sites [10]; therefore, this is where many parents are spending time and getting information. Moreover, 14% of all American mothers with at least one child in their household report turning to blogs for advice [11]. According to eMarketer, parenting Web sites are the top source moms use to learn about products and services [12], but mommy blogs and Web sites may also serve as sources of social support, connection, and validation for women navigating important health decisions for themselves and their family [13–15].

Digital and Social Media Utility and Limitations

The dialogic nature of social media allows senders to reach broad audiences and receivers to get involved in the conversation. A few studies have examined health promotion through social media [16, 17], the findings from which indicate potential for using blogs, Twitter and other online communication channels, not only for increasing awareness but also to influence decision making.

Specifically, there are a variety of digital and social media tactics that can be utilized for recruitment, for example paid advertisements on Facebook, Twitter, and Google and nonpaid outreach to parenting message boards and mommy bloggers; and these media offer advanced targeting capabilities, for example interests, audience, and geographic location, and real-time performance data analysis. Despite this, digital and social media also have limitations. For instance, research suggests that 90% of content is created by only 10% of users [18, 19], thus digital may not be representative and generalizability can be weak.

Use of Social Media in Research

Given these limitations with use of social media—and to inform the NCS social media feasibility activity—a meta analysis of more than 200 peer-reviewed and nonpeer-reviewed papers published between 2009 and 2013 was conducted to assess use of social media to support research. Findings included overall limited use with most of the publications focused on study recruitment and some focused on retention.

Among the recruitment-focused literature, differences between recruitment within patient-focused communities and more general

communities like Facebook and Twitter have been noted. Primarily, the differences center on the concept of the "ePatient," or the patient who actively seeks health information online, which the literature suggests is a different audience than those that might participate in other online communities such as Facebook or Twitter, and may be more easily engaged in health-related studies [20]. ePatients are 29% more likely to go to the doctor for regular check-ups and 21% more likely to be the first to try advanced medicines [20]. Findings from this area of work include successful use of social network recruitment for an HIV prevention education study [21]; a multimode recruitment method study which showed that internet advertisements yielded the largest number of recruited participants and completed surveys overall, yet Craigslist and email were more cost effective and successful at targeting young adult smokers who went on to complete the survey [22].

Comparatively, few articles focused on social media and retention. Despite this, the articles that did address use of social media for study retention suggest that social media platforms like Facebook can, in fact, help decrease attrition; encourage participants to engage in follow-up protocols; and generate active dialog, especially in the context of private user groups [3, 4]. Findings from this area of work include one longitudinal study that was able to locate 19 participants who would have been otherwise lost to the study—thus decreasing attrition by 16% [23]; and another study used Facebook to find adolescent girls who participated in an earlier study and recruit them for a follow-up study [24]. Out of the 175 girls, 78 were found on Facebook, 68 responded to the friend request, and 43 participated in the follow-up study [24]. Results showed promise for recruitment and retention of participants for studies on Facebook [24].

Discussion

Targeted solely to NCS participants in the western region of the United States, which was comprised of 10 states on or near the western coast, a private, closed social media-like community was developed. The goal of developing this private, closed social media-like community was to protect the privacy and identification of study participants; however, it was also grounded in social media best practices, such as development of social media appropriate content, open and transparent conversation, and active participation. The platform was built and tested in a user-centered design format, key features of which were to connect and share with other users; personalize one's community profile; and discuss topics and content and ask questions openly.

There are several implications researchers struggling with how to use social media effectively in their work that can be gleaned from this experiment. The following issues are discussed in the sections below: (1) sampling, (2) community design, (3) privacy and security, (4) Institutional Review Board (IRB) clearance, (5) community management, and (6) content development.

Geo-Based Cluster Sampling

To assess the impact, an experimental design was imposed. Participants for the private online community (the test group) were randomly selected from among participants. The balance of participants defined the control group. However, determining the most appropriate sampling strategy involved balancing efficiency with methodological rigor. The most efficient design would have been a balanced sample from across all western region study participants; however, this design posed a challenge. With the clustered nature of households in the sample segments, it was possible that participants who lived in close proximity to one another could reveal to others their invitation to join the private online community and this might have contaminated the control group by upsetting participants not selected to take part in the private online community (and

possibly driving them to not participate in NCS data collection activities entirely). One option in response to the contamination concern could have been to sample by Primary Sampling Unit (PSU). The western region was comprised of 10 PSUs and participants were assigned to the PSU nearest their home. However, sampling just by PSU could have created unintended sampling biases because of the similarities in culture, community type, and language among members of each PSU, which varied dramatically across PSUs.

Therefore, a geo-based cluster sampling approach was used to develop the sample frame. Roughly, the original sample segments were used to define the clusters of households invited into the private online community. For this, geographic information systems were used to map and group participants based on participants' geographic proximity to their PSU. The mean cluster size was 7.7, the minimum cluster size was 4, and the maximum cluster size was 16. After the clusters were formed, they were divided into 2 groups—a test group and a control group for each of the defined PSU regions. Using this model, 787 out of 1526 participants were invited to join the private online community. Fig. 1 displays the distribution of the participants initially assigned to the western region showing the control and test group.

Private Online Community Design

The private, invitation-only online community was conceptualized and designed to serve parents of young children. Data show that parents are heavy users of online social media and enjoy connecting online with other parents of young children about a variety of topics. One study found that, "75% of [parents] turn to social media for parenting-related information and social support" [25]. Yet, as discussed, it also had to follow privacy and security protocols set by the NIH. To accomplish both of these goals, the private online community was designed similarly to other popular online social networks that this group of users would be familiar with, for example, Facebook; yet instituted adjustments to maintain users' privacy.

The look and feel of the private online community was friendly, upbeat, and uncluttered. It allowed for personalization, but rather than having members use their own photo and name, they were prompted to choose a profile image, or avatar, from a selection of colorfully illustrated images and create a playful username from predetermined sets of nouns and adjectives to protect their identities. The private online community had a terms of service agreement that users had to accept before joining. This agreement outlined what content would be allowed and what content would be removed if posted. Acceptance of this agreement also acted as the consent process for participants to engage in the community, per discussions and approval of The *Eunice Kennedy Shriver* National Institute of NICHD IRB.

Security and Privacy

Because of the unique nature of this pilot—the use of social media (which is traditionally an open medium for discussion) among a group of study participants (for whom privacy and security are of the utmost importance)—steps had to be taken to ensure that any private online community developed for the NCS would meet the security protocols required of a federally funded research study. This meant participants' information needed to remain private and anonymous. To address these concerns, a privacy impact assessment was filed with NICHD, and system categorization documentation was filed with the National Institute of Standards and Technology. A vulnerability scan was also performed on the development version of the private online community to test for any security holes that could allow unauthorized access to the Web site. Ultimately, the pilot was classified as a low system security information system because participant personal information was not shared within the private online community.
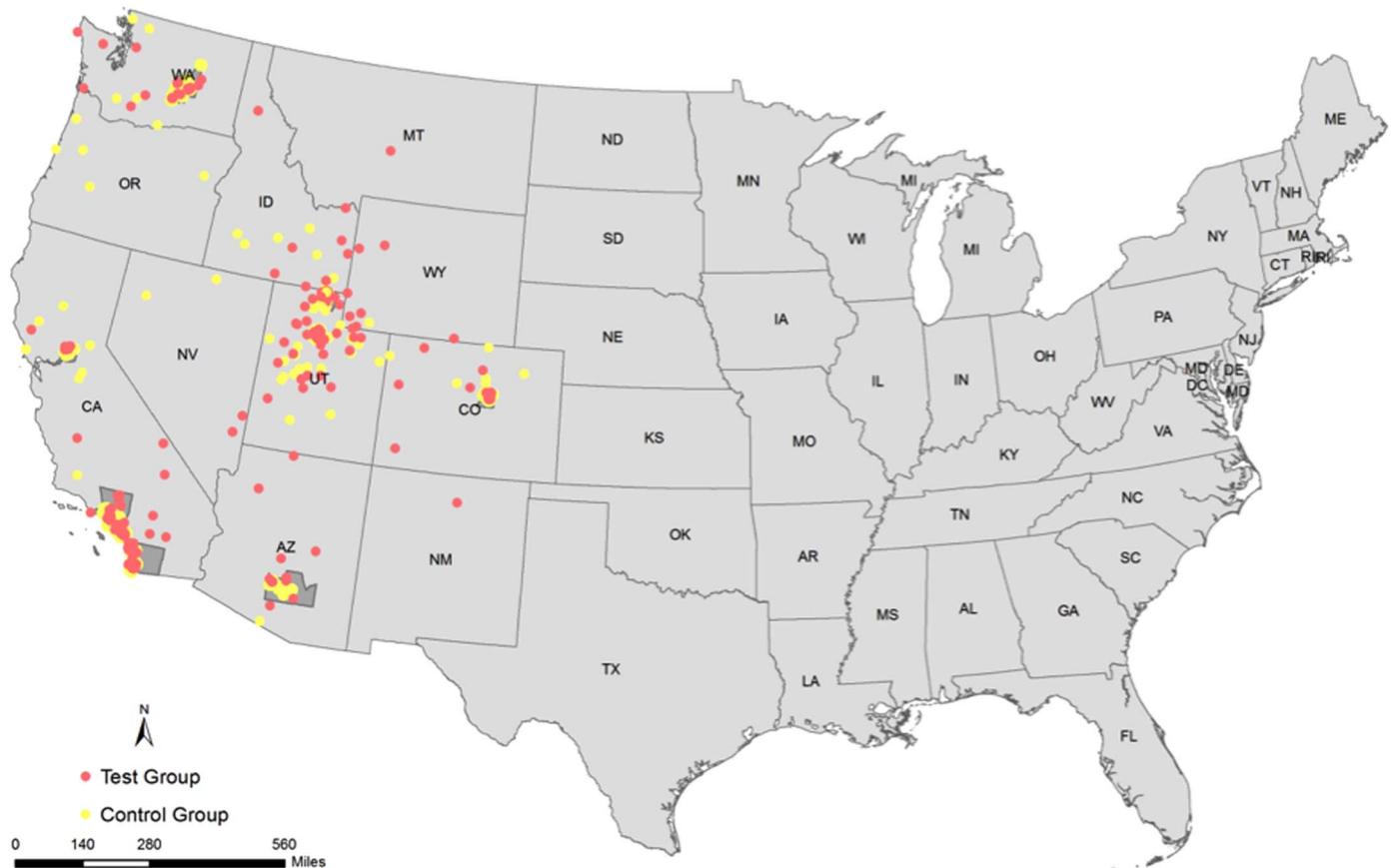
Fig. 1. Distribution of the participants initially assigned to the western region showing the control and test group. To protect the confidentiality of the participants, the data has been masked by random perturbation within a circle. This technique randomly moved the participant in any direction within a 100-mile radius.

IRB Clearance

The NCS posed new challenges for its IRB. IRB prospective clearance and the nature of social media are at odds with one another—social media, meant to be spontaneous and facile, is not a model of communications that IRBs are used to reviewing. For instance, at the time, there was no mention of social media in the common rule and the Office for Human Research Protections, and HHS had provided minimal guidance on communications with research participants through social media. Thus, the NCS relied on available information, such as other publications explaining the use of social media in research studies and HHS and NIH policies on the use of social media to develop guidance and ensure appropriate protections were in place.

The NCS went through several steps to explain the need for social media; how these media would be used; and how privacy would be maintained. The first interaction with the IRB was an overview presentation on social media and its importance to the NCS. Following the presentation, a detailed social media plan was developed to clearly outline the privacy protections being put in place. This plan included an editorial calendar, which mapped out the content of each post over the course of the pilot; a sampling plan, which included details on which participants would be asked to join the social media platform and which would serve as the control group; and a decision tree, which was developed to explain the decision process for addressing different interactions that took place in social media. For instance, it outlined anticipated comments and how the team would deal or respond to them. Finally, an electronic consent form was proposed to consent participating individuals. This was developed based on other applications approved by the IRB.

In terms of the private, closed online community, the IRB's questions focused on how to determine which participants would be able to join the platform; what type of content would be posted for them to read

and respond to; how would inappropriate content or self-disclosures be dealt with; and how would participants maintain privacy with individual profiles.

Clearance was obtained after responses to the questions were provided and the NCS agreed to provide the IRB with periodic updates and alert them if any issues arose. The NCS also worked with its IRB to clear all study content. This was done by submitting content in batches every 6 months.

Content Development

The team developed content in 6-month blocks (the second 6-month block was never cleared because the study was closed before it could be submitted). This content focused on topics of interest to parents of young children including child health, child development, nutrition and physical activity, and lifestyle. Content often linked to outside sources of information, drawing heavily on government sources to ensure the accuracy of information. This content was developed before launch and approved by the IRB. Once live, the community manager posted this preapproved content daily to encourage participant conversation.

Community Management

To ensure that the participants understood the purpose of the private online community and any protocols that had to be adhered to, an escalation and communications policy for identifying and responding to private online community posts was developed. This policy asked that participants refrain from sharing their names, locations, or other identifying information in the private online community. The policy was written in plain language to ensure comprehension, outlined policies to protect privacy, and provided parameters for what kind of information could be shared. This policy guided the daily monitoring of

the private online community and responses to participant inquiries. The community manager monitored comments and questions daily, responding as necessary to increase engagement.

Conclusion

Online social networks continue to be a part of the daily lives of Americans—and there is evidence that these online communities can provide a sense of belonging and support to their members. As it relates to use in longitudinal research, the opportunity to engage participants in an ongoing manner in this way can be enticing. However, these online communities are not without their challenges. Attention should be paid to balancing the openness of these kinds of communities with the privacy requirements of a research study. Consideration should also be given to the benefit that the study will gain by having the private online community available to all participants Versus the benefit of designing an experiment that can provide data showing effectiveness with a subpopulation. Planning and getting requisite approvals in a federal research environment, such an initiative as part of a research study can be a complicated process; more research is needed to look at the feasibility of using these online communities in an ongoing manner in research.

Acknowledgments

Declaration of Interest

The authors have no conflicts of interest to declare.

References

1. **Pew Research Center.** Social media fact sheet [Internet], 2017 [cited Sep 13, 2017]. (http://www.pewinternet.org/fact-sheet/social-media/)
2. **Pew Research Center Internet Science Tech RSS.** The social life of health information, 2011 [Internet], 2011 [cited Sep 13, 2017]. (http://www.pewinternet.org/2011/05/12/the-social-life-of-health-information-2011/)
3. **Rizvi SL, et al.** A pilot study of the DBT coach: an interactive mobile phone application for individuals with borderline personality disorder and substance use disorder. *Behavior Therapy* 2011; **42**: 589–600.
4. **Bolanos F, et al.** Using Facebook to maximize follow-up response rates in a longitudinal study of adults who use methamphetamine. *Substance Abuse: Research and Treatment* 2012; **6**: 1–11.
5. **Institute of Medicine (IOM).** Communicating with today's WIC mom: the millennial generation [Internet], 2011 [cited Sep 13, 2017]. (https://archive.cdph.ca.gov/programs/wicworks/Documents/Millennial%20Generation/WIConnects%20Presentations/Communicating_with_Todays_WIC_Mom_Meredith_10.3.11.pdf)
6. **Perrin A.** Social media usage: 2005–2015. Pew Research Center [Internet], 2015 [cited Sep 13, 2017]. (http://www.pewinternet.org/2015/10/08/social-networking-usage-2005-2015/)
7. **Moorhead SA, et al.** A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. *Journal of Medical Internet Research* 2013; **15**: e85.
8. **eMarketer.** What health info do consumers seek online? [Internet], 2013 [cited Sep 13, 2017]. (http://www.emarketer.com/Article/What-Health-Info-Do-Consumers-Seek-Online/1009698)
9. **Bailey M.** New survey reveals moms' media habits. Marketing to Moms Coalition and Current Lifestyle Marketing, 2008.
10. **Duggan M, et al.** Parents and social media. The Pew Research Center [Internet], 2015 [cited Sep 13, 2017]. (http://www.pewinternet.org/2015/07/16/parents-and-social-media/)
11. **Scarborough Research.** Mommy Blogger Infographic [Internet], 2012 [cited Sep 13, 2017]. (http://mashable.com/2012/05/08/mommy-blogger-infographic/)
12. **eMarketer.** Understanding how new moms share [Internet], 2010 [cited Sep 13, 2017]. (http://www.emarketer.com/Article.aspx?R=1008056)
13. **Heisler JM, Ellis JB.** Motherhood and the construction of "mommy identity": messages about motherhood and face negotiation. *Communication Quarterly* 2008; **56**: 445–467.
14. **Mitchell W, Green E.** 'I don't know what I'd do without our Mam' motherhood, identity and support networks. *The Sociological Review* 2002; **50**: 1–22.
15. **Smith JA.** Identity development during the transition to motherhood: an interpretive phenomenological analysis. *Journal of Reproductive and Infant Psychology* 1999; **17**: 281–299.
16. **Gustafson DL, Woodworth CF.** Methodological and ethical issues in research using social media: a metamethod of Human Papillomavirus vaccine studies. *BMC Medical Research Methodology* 2014; **14**: 1–11.
17. **Zhang C, Gotsis M, Jordan-Marsh M.** Social media microblogs as an HPV vaccination forum. *Human Vaccine and Immunotherapeutics* 2013; **9**: 2483–2489.
18. **Arthur C.** What is the 1% rule? *The Guardian* [Internet], 2013 [cited Sep 13, 2017]. (http://www.theguardian.com/technology/2006/jul/20/guardianweeklytechnologysection2)
19. **Nielson J.** Participation inequality: lurkers vs contributors in internet communities [Internet], 2013 [cited Sep 13, 2017]. (http://www.nngroup.com/articles/participation-inequality/)
20. **Orizio G, et al.** The world of e-patients: a content analysis of online social networks focusing on diseases. *Telemedicine and e-Health* 2010; **16**: 1060–1066.
21. **Rice E, et al.** Mobilizing homeless youth for HIV prevention: a social network analysis of the acceptability of a face-to-face and online social networking intervention. *Health Education Research* 2012; **27**: 226–236.
22. **Ramo DE, Hall SM, Prochaska JJ.** Reaching young adult smokers through the internet: comparison of three recruitment mechanisms. *Nicotine & Tobacco Research* 2010; **12**: 768–775.
23. **Mychasiuk R, Benzies K.** Facebook: an effective tool for participant retention in longitudinal research. *Child: Care, Health and Development* 2012; **38**: 753–756.
24. **Jones L, et al.** Recruiting adolescent girls into a follow-up study: benefits of using a social networking website. *Contemporary Clinical Trials* 2012; **33**: 268–272.
25. **Duggan M, et al.** Parents and social media. *The Pew Research Center* [Internet], 2015 [cited Sep 13, 2017]. (http://www.pewinternet.org/2015/07/16/parents-and-social-media/)