Supporting Statement B For:

The National Cancer Institute (NCI) SmokefreeTXT (Text Message) Program Evaluation

(3-21-2013)

Erik Augustson, PhD, MPH,

Behavioral Scientist/Health Science Administrator

Division of Cancer Control and Population Sciences

National Cancer Institute

6130 Executive Blvd, EPN-4034

Bethesda, MD 20892-7337

Telephone: 301-435-7610

Fax: 301-496-8675

E-mail: augustse@mail.nih.gov

Table of Contents

B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS.....1

B.1 RESPONDENT UNIVERSE AND SAMPLING METHODS.....1

B.2. PROCEDURES FOR THE COLLECTION OF INFORMATION.....4

B.3 METHODS TO MAXIMIZE RESPONSE RATES AND DEAL WITH NONRESPONSE.....7

B.4 TEST OF PROCEDURES OR METHODS TO BE UNDERTAKEN.....9

B.5 INDIVIDUALS CONSULTED ON STATISTICAL ASPECTS AND INDIVIDUALS COLLECTING AND/OR ANALYZING DATA.....**10**

LIST OF ATTACHMENTS

1. Evaluation Planning Matrix
2. Evaluation Conceptual Framework
3. Antifraud Measures
4. Literature Review Reference List
5. Data Collection Plan
6. List of Collaborators and Consultants
7. Online Consent Form
8. Institutional Review Board (IRB) Review
9. Screener Questionnaire*
10. Baseline Questionnaire*
11. Recruitment Outreach Plan
12. Honesty Pledge
13. Exit Scripts and Survey for Ineligibles*
14. 1 Week Post Quit Date Questionnaire*
15. 6 Weeks Post Quit Date Questionnaire*
16. Email and Phone Verification
17. Privacy Impact Assessment (PIA)
18. NIH Privacy Act Memo
19. 12 Week Post Quit Date Questionnaire*
20. 24 Week Post Quit Date Questionnaire*

* Indicates a survey.

B. STATISTICAL METHODS

B.1 Respondent Universe and Sampling Methods

For this study, we will recruit smokers aged 18 to 29 who have smoked on at least 5 of the last 30 days and who are interested in stopping smoking in the next 30 days. Participants will be recruited via e-mail outreach and online advertising to recruit smokers interested in quitting. MMG will develop and place ads on Facebook, Craigslist, and Pandora and through Google Display Network and Yahoo! Display Network (**Attachment 11**).

When online users click on the ad, a unique identifier will be created, which will allow us to link screener data and any survey data collected throughout the study. After clicking the ad, users will be directed to the study landing page and screener. Participants interested in the study will be screened to determine if they meet the eligibility criteria (**Attachment 9**). We will exclude those who are currently involved in a quit smoking program. Other eligibility criteria, including the technology requirements, are depicted in Figure 1.

Those who complete the screener and who are eligible for the study will be asked to complete an online informed consent (**Attachment 7**) and honesty pledge (**Attachment 12**). Once completed, participants will be directed to the baseline survey, randomized and enrolled in SmokefreeTXT.

Those who are deemed ineligible as a result of their answers to the screener will complete an exit survey (**Attachment 13**), containing a few demographic questions. These questions will allow us to compare those who were ineligible to participate in the study to those who are eligible and go on to complete the baseline survey, which includes the same demographic measures.

Figure 1. Eligibility Criteria for Study

Characteristic	Criteria
Age	18 to 29 years of age
Smoking Behavior	Smoked at least 5 days in the last 30 days
Interest in Smoking Cessation	Must be at least "mildly interested" in quitting smoking
Engagement in other quit smoking programs	Not using another program
Ability to Access Intervention	Have cell phone that has texting capabilities and willing to share phone number
Ability to Participate in Evaluation	Able to receive unlimited text messages on cell phone or willing to receive up to 130 text messages
	Have and are willing to share an e-mail address that can be used to deliver e-mail reminders/links to surveys and agrees to use it while participating in the study.
	Responds to confirmation email indicating email address is functional
	Mobile Commons confirms that phone can receive text message
	Willing to complete five surveys

Arm 1	Assessment-only Messages	Brief, periodic messages prior to participants' quit date and weekly assessment messages
Arm 2	Assessment + Countdown Messages	Weekly countdown messages prior to participants' quit date, and weekly assessment
Arm 3	Daily messages	The full Smokefree TXT program including daily messages comprised of countdown messages, assessment messages, tips and motivational messages.

The evaluation will take the form of a three-arm randomized trial with 24 weeks of follow-up (data collected at baseline and four follow-up periods). We will strive to enroll a total of 4,248 participants to complete the baseline survey over two years. Those who are eligible will be consented, enrolled, and then randomized to one of three arms:

If the participants are eligible they will complete the baseline survey (**Attachment 10**).

At the end of the survey, the participant will be assigned a quit date two weeks from completion of the baseline survey. To obtain feedback during and immediately following the intervention period, we will survey participants again at 1 week and 6 weeks after their quit date (which represents the end of treatment) (**Attachments 14 and 15**). To examine the longer-term outcomes of the interventions, we will survey participants at 12- and 24-weeks post-treatment (Figure 2). And to determine the degree to which participants engage with the texting program, we will use data about the text messages participants send to the program that can be linked to an individual study participant.

When determining the number of participants to recruit to the study, we estimated attrition across the data collection waves (Figure 1). Based on power calculations, we have set the goal for recruitment to be 4,248 smokers so that a small effect size could be detected at the 24-week post-treatment survey point (calculations assume 80% power and an alpha of .05). Figure 1 outlines our sample size calculations, illustrating the expected attrition across the waves of data collection and the implications on effect size.

Figure 2. Estimated Effect Size

	Interview Waves (5)				
		Post Treatment			
	Baseline	**1 Week Post Quit Date**	**6 Weeks Post Quit**	**12 Weeks Post Treatment**	**24 Weeks Post Treatment**
N per group	4,248	3,398	2,719	2,175	1,740
% attrition	N/A	20%	20%	20%	20%
Detectable effect size (alpha =.05, power = 0.80; 2-tailed test)	0.11 (small)	0.12 (small)	0.13 (small)	0.15 (small)	0.19 (small)

B.2 Procedures for the Collection of Information

Enrollment Process

After being screened into the study, participants will receive an online consent form (**Attachment 7**). Those who agree to participate will then be asked to acknowledge an honesty pledge (**Attachment 12**). Those who agree with the pledge will be asked to provide their contact information and gift card preference. They will then have to check their e-mail and click on a link to confirm they have a working email address. Mobile Commons will notify RTI if the phone number cannot receive text messages. Participants will be randomized to one of the three conditions. They will then complete the baseline survey (**Attachment 10**), which will include information to enroll them in the appropriate version of SmokefreeTXT.

Surveys

The five surveys will be administered via an online system, called Hatteras. After they have been screened and consented, have acknowledged the honesty pledge, and have provided contact information, participants will be directed to the baseline survey (**Attachment 10**). Seven days after their quit date, they will receive an e-mail invitation (**Attachment 16**) to take the first follow-up survey. This invitation will include a username and password, which they will need to log in to the follow-up surveys. At each follow-up period, participants will receive an e-mail inviting them to answer an online survey. A link to each survey will be provided in the e-mail. The baseline survey will assess some basic demographic and smoking history information and will include more items than the follow-up. Specific measures will include smoking status and smoking history and engagement with the website and/or text messaging and outcomes, including smoking cessation, reduction in number of cigarettes, self-efficacy regarding smoking,

skills building, and other proximal outcomes. Specific measures are presented in the Evaluation Planning Matrix (**Attachment 1**).

Participants' will be assigned a quit date approximately 2 weeks after the baseline survey is completed. This quit date will be used to calculate when e-mail invitations to complete each of the follow-up surveys should be sent. For example, the first follow-up survey will be sent 7 days after a participants' quit date. Nonresponse follow-up will be triggered when a participant has not completed the survey in a specified period of time. The SFTXT Evaluation Data Collection Plan (**Attachment5**) describes the protocol for handling e-mail (e.g., number of days after survey distribution date) and telephone reminders (e.g., number of call attempts, when to leave a message). Three additional follow-up surveys will be sent at 6, 12 and 24 weeks.

B.2.1. Quality Control

The contractor for this study will establish and maintain quality control procedures to ensure standardization, and high standards of data collection and data processing. An important first step in the analysis phase of a study is to ensure the quality of the data. Quality can be defined in multiple ways, and therefore, we define the checks and data editing procedures below.

(1) *Skip patterns*. Some questions in the instrument may be skipped entirely given responses to a prior question. Analysis of internal testing data and field test responses will provide a verification that the web-based questionnaire skip patterns are working correctly.

(2) *Consistency checks*. Once the questionnaires have been finalized, we will develop a detailed data editing plan to identify questions that should have consistent responses (e.g., time till first smoke in the morning and least preferred time of the day to not smoke), along with the data editing rules for dealing with any inconsistencies. They may include

deletion of one or more inconsistent values, imputation of inconsistent value(s), and no change.

(3) *Partial versus complete interviews*. Some sample members will complete the full web-based interview, while others will stop the interview midprocess never to return. The decision rule will be specified in the data editing plan to categorize sample members at each wave as respondents, partial respondents (sufficient responses for analysis), and nonrespondents (complete nonresponse or insufficient responses for analysis).

(4) *Item nonresponse*. The distribution of question responses, including item nonresponse, will be summarized in a codebook at each wave of the study to accompany the SAS file. Variables with high levels of nonresponse, such as 50% or more, will be identified and marked for potential exclusion from any subsequent analyses.

(5) *Imputation*. Missing values, as noted for nonresponse bias, are problematic in that they can reduce the power of statistical models. Variables important to the analysis will be evaluated for imputation by examining the patterns of (item) nonresponse within a round of the study in addition to patterns of (unit) nonresponse across the waves. The pattern of missing data may dictate the type of methodology used such as (sequential) hot-deck or a model-based approach. If used, comparisons of data distributions before and after imputation will be conducted to ensure that the procedure does not introduce unnatural patterns in the database. In addition, imputation flag (indicator) variables will be included on the SAS file for easy identification.

(6) *Derived variables*. Additional variables may be required for the analyses that are derived from responses obtained from a particular study wave or composited across waves of the study. These variables will be identified upon approval of the analysis and documented

for the final report. Quality checks will be implemented to ensure the accuracy of the derived variables. Variable labels on the SAS data file will identify the information as being derived.

The contractor will maintain a log of all decisions that affect sample enrollment and data collection. The contractor will monitor response rates and completeness of acquired data.

B.3 Methods to Maximize Response Rates and Deal with Nonresponse

Understanding the potential impact of nonresponse bias will be important. If everyone for whom one or the other intervention is not helpful drops out and fails to complete surveys beyond the baseline, then nonresponse bias may unfairly lead to a higher proportion of cessation success than is warranted (as data will come from remaining respondents). Theoretically, respondent attrition is only a major threat if such attrition differs substantially between the two conditions. Nonetheless, we will compare those who complete all four waves of interviews against those who drop out after the baseline survey to assess whether there are major demographic or other differences. Understanding the potential impact of nonresponse bias is important for correctly interpreting the study results. Nonresponse bias can occur in three ways as discussed below.

Unit nonresponse to the baseline survey. It is the hope of researchers using online panels to recruit participants that collectively represent the population of interest. Otherwise, the generalizability of the results comes into question. For example, if all study participants were from high socioeconomic status (SES), then the study results may pertain only to a similar population if high SES was linked to smoking cessation/continuation. Consequently, without a sampling frame containing population information, nonprobability samples must be compared

with extant data sources typically. RTI will compare the distribution of characteristics for the study enrollees against demographic information provided by such sources as the Decennial Census (www.census.gov). Excessive levels of nonresponse bias at baseline will initiate an evaluation of analysis weights for the study.

Study withdrawal (unit nonresponse) after the baseline interview. By comparison, the baseline survey provides a wealth of information for evaluating nonresponse bias in later study waves. The importance of assessing nonresponse at later waves is demonstrated through an example. If everyone for whom their respective intervention is not helpful drops out and fails to complete surveys beyond the baseline, then nonresponse bias may unfairly lead to a higher proportion of cessation success than is warranted (as data will come from remaining respondents). Theoretically, respondent attrition is only a major threat if such attrition differs substantially between the two conditions. Nonetheless, we will compare those who complete the interviews against those who drop out after the baseline survey to assess whether there are major demographic or other differences.

Nonresponse bias is estimated by

$$Bias\left(\hat{\theta}_R\right) = \hat{\theta}_R - \hat{\theta} \qquad (5.1)$$

where $\hat{\theta}_R$ is the statistic estimated from the survey respondents to any particular wave (or all waves), and $\hat{\theta}$ is the corresponding estimate using all sample members regardless of response pattern. Note that the study information used to generate the statistics must be known for later-wave respondents and nonrespondents (i.e., baseline values).

Item nonresponse. Nonresponse bias can also occur if differential patterns of missing question item responses exist in the data for the participating sample member. Similar to unit nonresponse, item nonresponse reduces the power of statistical models when the desired

covariates have high levels of missing values. We will conduct an item nonresponse analysis using expression (5.1) for a set of key items known for respondents and nonrespondents where response is defined on an item-by-item basis.

Beyond the primary outcome of sustained smoking avoidance, we can attempt to determine whether intervention components affected more proximal outcomes, including self-efficacy, skills building, and motivational salience. In addition to testing for main effect differences between the three experimental conditions, we could conduct subgroup analyses or explore interaction effects to determine whether the intervention was more effective for particular demographic groups or whether other factors, such as psychosocial variables or smoking history, moderated the effects. We could potentially also consider path analysis to test hypotheses related to mediation of the effects. For example, we can look at whether the effects of the intervention were mediated through self-efficacy, skills building, or motivational salience.

B.4 Test of Procedures or Methods to be Undertaken

The majority of items used with all surveys were derived from existing surveys, including the Youth Risk Behavior Survey (YRBS), Youth Tobacco Survey (YTS), Tobacco Use Special Cessation Supplement to the Current Population Survey (TUS-CPS) to identify tobacco use and cessation-related measures, the Health Information National Trends Survey (HINTS) and the Pew Internet and American Life Survey. Additionally, delivery of the text messages and user responses has already been successfully tested in the established SmokefreeTXT program. Finally, a two week pre-testing period will be performed to ensure quality control in which the first respondents will be part of the pre-testing period.

Reviewing how well processes are working in the first few weeks of the study will allow us to make any necessary adjustments. In addition, reviewing data gathered in the first two weeks of the study will also help identify any issues with item non-response or lack of variability in responses. Should there be any non-substantive changes to the methodology or instruments, a change request will be submitted to OMB for review and approval.

B.5 Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

There were many individuals critical in developing the research plan, the conceptual framework, survey questions, and sampling strategies underlying the SmokefreeTXT Evaluation. Many of the same individuals will be involved with analysis once the data are collected, and are listed in **Attachment 6**.