

Leveraging the emerging field of disaster citizen science to enhance community resilience to improve disaster response

Supporting Statement – Section B

Submitted: 02/21/2018

Project Manager

Mary Leinhos, PhD, MS

Office of Science and Public Health Practice

Office of Public Health Preparedness and Response (OPHPR)

2877 Brandywine Rd., Mailstop K-72

Atlanta, GA 30341

Telephone: (770) 488-8619

Fax: (770) 488-8688

Email: MLEinhos@cdc.gov

Table of Contents

1. Respondent Universe and Sampling Methods
2. Procedures for the Collection of Information
3. Methods to Maximize Response Rates and Deal with No Response
4. Test of Procedures or Methods to be Undertaken
5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

List of Attachments

- Att. A Authorizing Legislation
- Att. B 60-Day FRN
- Att. C Consent Forms
- Att. D. Interview Guide
- Att. E Survey
- Att. F Recruitment and Reminder emails
- Att. G RAND IRB Determination
- Att. H CDC Human Subjects Determination

SUPPORTING STATEMENT B

Leveraging the emerging field of disaster citizen science to enhance community resilience to improve disaster response

1. Respondent Universe and Sampling Methods

The Centers for Disease Control and Prevention (CDC) Office of Public Health Preparedness and Response (OPHPR) proposes a new information collection to learn about how the emerging field of disaster citizen science can enhance community resilience to improve disaster response. This information collection will explore the potential of disaster citizen science, including facilitators and barriers, for increasing community resilience, enhancing participation in preparedness and response activities, and improving preparedness efforts. Collecting data on this topic will inform the development of educational and instructional tools for communities and health departments to navigate the emerging field of disaster citizen science and promote collaborations. Insights gained from this information collection will assist in strengthening the evidence-base on disaster citizen science, including when and how it can be used, and how to structure citizen science projects for greater benefits and positive impacts.

Information on disaster citizen science will be obtained through 35-55 individual and group interviews with citizen science stakeholders including citizen scientists and public health officials in a position to use citizen science data to inform decision-making. It will also be obtained through a survey fielded to 600 local health departments (LHDs) from across the U.S.

Individual and group interviews: Interviews will be conducted with stakeholders located across the U.S. Participants will be selected among citizen science organizations and their partners (e.g., academics who conduct research on citizen science) and public health organizations that have used or considered using citizen science data to inform decisions. The researchers will sample for maximum variation, seeking to obtain variation on U.S. region, type and sophistication of citizen science project, for diversity in representation of communities (e.g., urbanicity, population size), type of disaster encountered, and previous experience with disaster citizen science. The contractor, RAND, will identify potential participants through literature reviews as well as snowball sampling in a phased approach that starts with citizen science and LHD organizations (e.g., National Association of County and City Health Officials (NACCHO) and the Citizen Science Alliance) with the broadest scope and ask for recommendations for participants that vary with respect to the aforementioned criteria.

Survey: For the survey, we seek to obtain a nationally representative sample and we will apply survey weights to ensure that findings have external validity and can be generalized to LHDs in the U.S. The sample for this survey will be drawn from an existing database of all LHDs (n=2,800), in the U.S., which is maintained by the National Association of County and City Health Officials (NACCHO). This database includes the name and contact information of a health official at each LHD as well as basic characteristics of each LHD the individual is affiliated with (e.g., size of population served, state). For the survey, we will follow validated methods used for over twenty years in the NACCHO National Profile of Local Health Departments Survey to produce national estimates of LHD characteristics and activities.

NACCHO will pull a stratified random sample of 600 local health officials representing diverse LHD. Strata will be defined by size of the population served by the LHD. We will use the following strata defined in the Profile survey: <25K; 25-45K; 50-99K; 100-249K; 250-499K; 500-999K; and 1,000,000+.¹

On average, the response rates for surveys of LHDs using this sampling frame range from 50-60%. Therefore, NACCHO will pull a sample of 600 LHDs in order to obtain 300 completed surveys. RAND will request available data on characteristics of non-responders. We are confident that this effort will yield a nationally representative sample. Because we are sampling from the full universe of health departments, we will investigate whether those who respond to the survey are systematically different from those who do not, and we can then adjust for any selection bias we can identify. Furthermore, we will use sampling weights to ensure representativeness.

The stratified random sample of 600 LHDs balances the need for adequate precision for population-level estimates while reducing unnecessary burden on local health officials. For example, we did not want to exceed a sample size requirement in excess of 25% of the population size. Using a hypothetical calculation with a prevalence estimate (p) of 50%, which yields the most conservative scenario, we found that a sample size of 600 would result in a margin of error of +/- 4%, or a 95% confidence interval of (46%, 54%) around p . As survey estimates approach 0% or 100% (and away from 50%), the precision improves. To evaluate the precision around subgroup differences, we assumed an equal allocation of 200 per stratum (based on three strata defined by size of population served- <45K, 50-499K, and 500K+). If we were to test for significant differences between prevalence estimates in any two strata (assuming Type 1 error rate of 5% and power of 80%), the smallest difference we can detect as statistically significant varies between 9% (for $p=10%$) and 13% (for $p=50%$). Note that a prevalence estimate of 50% yields the worst precision, and we anticipate few of our estimates to be at 50%.

2. Procedures for the Collection of Information

The information collection tools (interview guide and survey) have been submitted as part of this information collection in the statement provided to OMB (see Attachment D Interview Guide and Attachment E Survey). Information collection tools have been reviewed and approved by the IRB at the contractor. CDC human subjects approval has been obtained. Approval letters from the contractor's IRB and CDC's human subjects determination are included with this new information collection request (see attachments to Supporting Statement A).

In this mixed-methods project, we will conduct qualitative and quantitative research to meet two objectives: (1) explore the potential of disaster citizen science for increasing community resilience, enhancing participation in preparedness and response activities, and improving preparedness efforts; and (2) develop educational and instructional tools for communities and health departments to navigate the emerging field of disaster citizen science and promote collaborations. First, we will conduct approximately 35-55 (groups and individual) interviews with two different populations of citizen science stakeholders: (1) citizen scientists and the organizations which represent them; and (2) public health officials in a position to use citizen

¹ http://nacchoprofilestudy.org/wp-content/uploads/2017/04/ProfileReport_Final3b.pdf

science data to inform decisions. In addition, we will field a nationally representative survey to officials in 300 LHDs and develop and pilot test citizen science toolkits with approximately three LHD-community group dyads.

Collection and analysis of qualitative interview data

For the interview component of the project, RAND will recruit a maximum of 35-55 interview participants from each of the following stakeholder populations: (1) citizen scientists previously and/or actively involved in data collection and their partners (e.g., academics who study citizen science and engage in citizen science projects); and (2) LHD health officials and emergency managers engaged in official preparedness and response activities. We will identify potential participants by searching online databases of citizen science projects that include contact information of project leaders and partners, Lexis Nexis searches (i.e., to identify individuals quoted in the media about citizen science), posting interview requests to various listservs monitored by citizen science stakeholders, and by snowball sampling. Potential interview participants will be invited to participate via email. Interview participants will be asked open-ended questions about citizen science projects and activities in their communities and their knowledge and attitudes about the wider field of citizen science. We will seek subjective insights in participants' own words.

The interviews will be semi-structured. While there will be an interview protocol for each stakeholder group to help spur discussion, interviewers will probe on interesting themes that surface and will be encouraged to deviate from the protocol to obtain rich descriptions of unanticipated themes. In addition, data analysis will occur on an ongoing basis, and we will ask interview participants to comment on themes that are emerging from the data in order to build on and refine our theory of citizen science as well as our sampling strategy. Each interview will be driven by the particular stakeholder group, the participant's unique background and role in citizen science, and our working theory or understanding of disaster citizen science. Participants will be expected to provide different levels of information that then can be analyzed to make connections among the data.

Collection and analysis of survey data

RAND will field the 30 minute survey to approximately 600 local health officials identified by NACCHO. Because of a track record of response rates of 50-60% using this sampling frame, this should be enough to obtain 300 completed surveys.

RAND will field the survey via SelectSurvey, a secure platform for the management of online surveys. RAND will send out email invitations to all 600 selected local health officials. In the e-mail, selected local health officials will be given the opportunity to designate another staff person at their LHD as the primary contact to complete the survey. The e-mail will include a link to a Web-based questionnaire (housed on SelectSurvey), individualized with preloaded identifying information specific to the LHD. Local health officials can print a hard copy version of their questionnaire by using a link in the introduction to the Web-based questionnaire or can request that RAND staff send a copy via e-mail or U.S. mail.

Items of information to be collected

Interviews: Data collection efforts will engage citizen scientists and their partners (e.g., academics who work with citizen scientists on research projects) and local health officials in a position to use citizen science data to inform public health decision-making. Below we have outlined the information that will be collected during interviews:

- Promising practices in citizen science
- Barriers and facilitators of disaster citizen science
- Benefits and uses of citizen science data in disasters
- Characteristics of successful citizen science projects
- Comparisons of citizen science research versus “traditional” research
- Established and emerging technologies and their potential to support citizen science efforts
- Quality, validity, and reliability of citizen science versus traditional data in disasters
- Impacts of citizen science efforts including unintended consequences of the generation and use of citizen science data
- Strengths and limitations of citizen science resources
- Theories and models of citizen science

Survey: Data collection efforts will engage LHDs. Below we have outlined the information that will be collected in the cross-sectional survey:

- Knowledge, awareness, and attitudes about citizen science
- Experiences with citizen science
- Potential applications of citizen science
- Barriers to citizen science
- Readiness for citizen science
- Resources required for citizen science
- Drivers of citizen science (e.g., community engagement, commitment to research)

Estimation procedures

Interviews: Interviews will be coded for themes using standard qualitative analysis techniques. Themes as well as illustrative quotes will be described in peer-reviewed articles. Articles will also include information on the most frequently mentioned applications of disaster citizen science (e.g., for mosquito monitoring, recovery needs assessment) and barriers and facilitators of citizen science. Much of what we learn (including best practices in citizen science for public health) will inform the development of tools and guidance for LHDs and community groups engaged in citizen science.

Survey: At survey close, NACCHO will compute sampling weights that account for non-response and oversampling of large LHDs, using data from the sample and the sampling frame. All analyses conducted by RAND will employ sampling weights. The use of probability sampling and sampling weights will enable RAND to generate estimates representative of all LHDs in the U.S. Since this study is the first of its kind and exploratory, the initial step will employ descriptive analyses. For example, we will generate nationally representative prevalence

estimates (means, percentages) of measures such as local health official awareness of, barriers to, and experience with citizen science activities and data. In addition, we will compute these estimates for subgroups of interest for which we have adequate sample size (and precision) of respondents (e.g., differences by LHD size, disaster risk/experience). If there is sufficient variation in the outcomes of interest, we will also fit separate logistic regression models by outcome to identify predictors of engagement with citizen science/organizations and use of citizen science data to inform decision-making.

Degree of accuracy needed for the purpose described in the justification

OPHPR collects information in order to develop guidance and tools for local health departments and community groups on the topic of disaster citizen science. This type of guidance will help stakeholders to align their efforts for greater impact. Because of the exploratory purpose of this study, its accuracy will depend on representation from diverse stakeholders across the U.S. and input provided to the discussion by each participant. These can be facilitated by careful selection and recruitment strategies and adequate training provided to interview moderators.

Unusual problems requiring specialized sampling procedures

Unusual problems requiring specialized sampling are not expected with this new information collection. If such situations occur during the course of the project implementation, requests for changes in the proposed methodologies will be provided to OMB.

Any use of periodic (less frequent than annual) data collection cycles to reduce burden

Interviews: To reduce burden, each respondent will participate in only one individual or group interview. No follow-up interaction with participants will be needed. Total estimated burden time for each interview is 75 minutes (60 minutes to participate in the interview and 15 minutes for scheduling).

Survey: To reduce burden, each respondent will be asked to complete one survey. No follow-up interaction with participants will be needed. Total estimated burden time for each survey is 30 minutes.

3. Methods to Maximize Response Rates and Deal with No Response

Interviews: The following are examples of procedures that have proven effective in previous studies and will be used when possible to obtain at least a 90% response rate:

- Informing respondents of project purpose and rationale of the study, who will see the study results, and how the results will be used
- Using culturally appropriate data collection instruments and procedures
- Using alternative communication means, such as video-or teleconferencing
- Addressing data security and anonymity with respondents
- Minimizing the time needed for participation in the project
- Informing respondents about the project process and interview duration and setting, so that

they know what to expect

- Limiting participation of each respondent to one interview meeting
- Discussing the importance of the study for public health decision making and how the findings will be put into action
- Giving participants multiple options for interview times and allowing them to select a time that accommodates their schedule

Survey: RAND will conduct a variety of efforts to maximize response rates. All of the methods described here have been used by NACCHO in the past to secure high response rates for its National Profile of Local Health Departments Survey. First, we will offer respondents flexibility with respect to survey mode. For example, although the default option is to complete the survey online via SelectSurvey, we will allow local health officials to print a hard copy version of their questionnaire by using a link in the introduction to the Web-based questionnaire and/or request that RAND staff send a copy via e-mail or U.S. mail. Second, we will conduct follow-up with non-responders using email messages and telephone calls. Non-responders will receive up to three follow-up emails and one phone call to encourage participation. The survey will be in the field for a period of three weeks, allowing local health officials ample time to conduct the survey on their own time during this period or to designate another staff person within the LHD to complete the survey on their behalf. Finally, NACCHO and RAND will offer technical support to survey respondents through an email-address and telephone hotline.

4. Test of Procedures or Methods to be Undertaken

RAND pre-tested the survey with five local health officials that represent the diversity of the full sample. RAND conducted interviews to assess whether certain sections and questions performed as expected, to identify any questions and wording that were unclear, assess whether response options were appropriate, and estimate the length of time to required to complete the survey. Multiple adjustments were made to survey length and content following pre-testing. The interview guide will not be pre-tested, as it is designed to be semi-structured.

The information collection tools that will be used in this project are linguistically and culturally appropriate for the targeted populations. The importance of utilizing culturally and linguistically appropriate instruments and procedures is well-documented in the literature and is an important aspect of designing and implementing OPHPR's activities and programs.

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

The survey, sampling approach for surveys and interviews, and data collection procedures were designed by the RAND Corporation under the leadership of:

Ramya Chari, Ph.D.
1200 South Hayes Street
Arlington, VA 22202
(703) 413-1100 x5216
rchari@rand.org

Lori Uscher-Pines, Ph.D.
1200 South Hayes Street
Arlington, VA 22202
(703) 413-1100 x5167
luscherp@rand.org

The contact for additional survey questions at NACCHO is:

Laura Biesiadecki, Ph.D.
Senior Director, Preparedness, Recovery, and Response
NACCHO | Washington | DC | 20036 | USA
Lbiesiadecki@naccho.org 202-507-4205