Form Approved

OMB No. 0990-0390

Expiration Date 04/30/2018

**The Healthy Behavior Data Challenge**

Phase 1 Submission Template

**Introduction**

The Healthy Behavior Data Challenge responds to the call for new ways to address the challenges and limitations of self-reported health surveillance information and tap into the potential of innovative data sources and alternative methodologies for public health surveillance.

The Healthy Behavior Data Challenge will support the development and implementation of prototypes to use these novel methodologies and data sources (e.g., wearable devices, mobile applications, and/or social media) to enhance traditional healthy behaviors surveillance systems in the areas of nutrition, physical activity, sedentary behaviors, and/or sleep among the US adult population aged 18 years and older.

The collection of health data through traditional surveillance modes including telephone and in-person interviewing is becoming increasingly challenging and costly with declines in participation and changes in personal communications. In addition, the self-reported nature of responses particularly in the areas of nutrition, physical activity, sedentary behaviors, and sleep has been a major limitation in these surveillance systems, since self-reported data are subject to under/over reporting and recall bias. Meanwhile, the advent of new technologies and data sources including wearable devices (Fitbit, Garmin, Adidas, Jawbone, smart watches, activity trackers, etc.), mobile health applications on smartphones or tablets, and data from social media represents an opportunity to enhance the ability to monitor health-related information and potentially adjust for methodological limitations in traditional self-reported data.

The Healthy Behavior Data Challenge will harness this potential and identify feasible alternative options for collecting health-related behaviors in new ways. Conducted in two phases, Phase I (Prototype Development) entails Challenge participants developing a concept proposal for obtaining data collected from wearable devices, mobile applications and/or social media for public health surveillance purposes.

The Healthy Behavior Data Challenge participants will propose data sources and approaches for aggregating data from wearable devices, mobile applications and/or social media in the areas of nutrition, physical activity, sedentary behaviors, and/or sleep. In Phase II (Prototype Implementation), a subset of submissions (up to 3) with promising concepts will be invited to test their proposed approaches for ongoing public health surveillance.

**Website**:

Additional Information:

Information on the Behavioral Risk Factor Surveillance System can be found at [www.cdc.gov/brfss](http://www.cdc.gov/brfss). Details on the HBD Challenge may be found at challenge.gov.

For Further Information Contact: Dr. Machell Town at BRFSSinnovations@cdc.gov.

**Submission Deadline**:

1. Challenge Team Information

|  |  |  |
| --- | --- | --- |
| Team Name |  |  |
| SocialETC | | |
| Team Lead |  | City/Province |
| George Gruse |  | Brookevile, MD |
| E-mail |  | Phone Number |
| georgegruse@gmail.com |  | 240-482-7016 |
| Subject-matter/domain expertise |  |  |
| Main Programmer |  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Team Member #1** |  | **E-mail** |  | **Subject-matter/domain expertise** |
| **Greg Gruse** |  | [gggruse@socialetc.com](mailto:gggruse@socialetc.com) |  | **Programming Consultant/ Adviser** |
| **Team Member #2** |  | **E-mail** |  | **Subject-matter/domain expertise** |
| **Gergory Gruse** |  |  |  | **Programming** |

|  |
| --- |
| **Are all team members residents of the United States?** |
| **Yes** |

1. Organization (if submitting on behalf or as part of an organization)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Organization Name** |  | **Website** |  | **Type of Organization** |
| SocialETC |  |  |  | **LLC** |

1. How did you find out about this challenge?

|  |
| --- |
| Challenges.gov |

1. Submission Overview

|  |
| --- |
| **Project Title** |
| Healthy Behavior Research App |
| **Project Overview** |
| Describe in 500 words or less:   * What aspects of sleep, physical activity, nutrtion, and sedentary behavior do you propose to report on and why are they important for public health surveillance? * Provide a brief description of the source(s) of data that will be used to report on these aspects, how your team proposes to access them, and why they are appropriate for use in public health surveillance? * How do you see your concept improving on current public health surveillance in the areas of sleep, physical activity, nutrition, and sedentary behaviors?   **The Healthy Behavior Research App a new way to track healthy behaviors. Leveraging new innovative activity planning, tracking screens, tasks and surveys to capture healthy behaviors from research participants and the general public. Our goal is to create a mobile application that is both a useful app to users wanting to lead a health lifestyle as well as a research app where they can have their usage data included in a very large scale research study. Healthy Behavior Research App relies on Apple;s Research kit, an Open Source Development Kit designed to capture user input and store it on the phone for later distribution to web based data solutions. This project, will record only the data the users have given  consent on. The studies the subject takes part in will be selectable IE:**   * **A subject looking to participate in Physical activity may want to participate in the nutrition study as well. However, they may not want or be as applicable to the Sedentary study as well.** * **Conversly a Subject in Sedentary study may choose that physical activity is important as well.**   **Any possible combination is possible or even all of them. Our goal is to give user control and keep the strain of a study to the minimum.**  **We plan to capture, data over the subject's participation over long periods of time. The main goal is to use the Apple watch and Apple Health Kit to capture workout and physical activities to measure accurate health data. For example on the press of the start button,**   1. **The duration of the exercise/work will be logged** 2. **The date/time will be stored for analysis across the day/week/month/year.** 3. **Heart rate will be gathered to help track how various factors impact their strenous activity.**     1. **Weight**    2. **Height**    3. **Age**    4. **Gender**    5. **Location**    6. **Work** 4. **We can measure seasonal activity and climate based information to help determine factors that encourage or discourage physical activity.** 5. **The temperature and general weather outside.** 6. **The response of the subjects to the weather conditions.**   **We intend to use this information to help the CDC ascertain reasons as to many subjects obesity issues. How, the environment around them impacts their decision making in accordance with responses they give and the aggregate of the data collected. This is just one of the 4 studies we will include in the app.**  **All the data will be stored on a secure database hosted by Sage Bionetworks (which is used for several other research app) for security and ease of access. The CDC will be given a key for simple access to read or write data if need be.**  **Note: In the Phase 2 protoype development we will only implement the Mobile Application, it will be hooked up to the Sage Bionetworks Bridge SDK kit but the hosting side of the application will not be developed until after Phase 2.** |

1. Indicators to be measured (the indicators listed below are not comprehensive and innovators are recommended to include other relevant indicators)

* 1. Physical Activity
* Amount of MVPA[[1]](#footnote-0) time per day
* Amount of MVPA time per day obtained in bouts of 10 minutes or more
* Frequency of MVPA
* Calories burned
* Type of activity (aerobic, strength, etc.)
* Level of activity (low, moderate, high)
* Location of MVPA (recreation facility, at home, at work, on sidewalk/bike lane)
* Perception of safety while active
* Enjoyment level of the MVPA
* Average and peak heart rate
* Hours per week adults spent in sports, fitness or recreational physical activities
* Other indicators
  1. Sleep:
* Hours of sleep per night (sleep duration)
* Amount of time awake (sleep quality)
* Number of times awake (sleep quality)
* Number of adults reporting having trouble getting to and staying asleep
* Heart rate
  1. Sedentary Behaviors:
  + Average number of hours per day spent sedentary, excluding sleep timeAmount of time awake (sleep quality)
  + Average number of hours per day spent on a computer/screen including watching TV, videos, playing computer games, emailing or using the internet
  + Average number of hours per day spent on a computer/screen including watching TV, videos, playing computer games, emailing or using the internet
  1. Nutrition:
  + Total calories consumed per day
  + Consumption of fruit (not including juices) by day, week, or month
  + Consumption of green leafy or lettuce salads, with or without other vegetables, by day, week, or month
  + Consumption of green leafy or lettuce salads, with or without other vegetables, by day, week, or month
  + Number of sugar-sweetened beverages consumed by day, week, or month
  + Number of caffeinated drinks consumed by day, week, or month

1. Summary of proposed data source(s) (complete applicable sections)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Data Source** | | **Data Accessibility** (e.g., API, specialized software, existing data set) | **Data Cost** (i.e., fee for access, open access) | **Data Recency and Update Frequency** (i.e., how recent is the data and how often is it collected) | **Applicable Functional Area(s) and Indicator** (i.e., physical activity, nutrition, sleep, and/or sedentary behavior) | **Existing Users of the Data Source** (i.e., identify examples of organizations or other groups that have or are using the data source) |
| **Organization (e.g., company)** | Method of Collection (e.g., wearable, self-reported) |
| *1* | Social Etc. | Apple Health Kit Extension | SDK | Free | On use | Physical Activty | None |
| *2* | Social Etc | Apple Research Kit App | SDK | Free | On use | Physical Activity | None |
| *3* | Cardiogram | Algorithim | Reaching out to them to capture abnormal heart rythms to detect heart disease. | unsure | It would be collected extremely frequently. | Health | Cardiogram |

1. Describe how the data that you will use provides information and insight that is complementary to or more novel and innovative than that currently utilized for public health surveillance by CDC? (Novelty/innovation can apply at the level of the individual data source(s) selected, the specific indicators to be measured, tools/solutions that are used to capture the data, or result from newly created linked data sets). (750-word limit)

The mass data collection of heart rate and exercise information has not been done on this large a scale. This is rarely done in the research subject natural environment or in a non cumbersome methodology. The ease of use and accessibility will allow for researchers to study how people naturally workout.

Healthy Behavior Research App physical activity study the data we collect to be collected largely from the single press of a button for high user retention. Encumbering the user with large menus that are largely unimportant to the user raise the risk of subjects dropping the app and leading to incomplete data.

As such we made the choice that the bulk of the data should be captured in a single button press. We found that we can aggregate enough data from that button press and automate the rest so the device takes care of the work for them. With the integration of a apple watch we can record their heart rate as well to gather more information. This can help track and determine if a user has an irregular heartbeat or associated Heart health problems. This app strives to use the simple one track feature to gather as much data as possible to guarantee our users health.

Monitoring how and why people work out can help lead to advances insights into possible reasons people give up or avoid physical exertion.

For sleep we hope to determine how the enviorment may be affecting the subjects sleep patterns. The application can be used as a bed side alarm clock enabling the application to track when the user goes to sleep and when they wake up in the morning. The Healthy Behavior Research App tracks information about how the subjects enviorment and the general trend on amount and quality of sleep. Questions and the subjects answers can help find aspects of the subject enviorment that they may be able to change to improve the quality of their sleep. Subjects, spouses are usually affected by the sleep condition of their partner. This can lead to stress and restlessness for them as well. Improving the subjects sleep should improve the quality of sleep for the spouses/partners. This study is more exploratory and research into it should yield unique results that are hard to predict off hand. However, as was stated this study simply will strive to improve the quality of sleep for particpants and their spouses.

For nutrition, our goal is to leverage the mobile application to help curve obesity and find good trends that public health services can help promote to curve obesity. Users will use the app to plan their diet and track their progress on the diet. This study in conjunction with the Physical activity study will see how we study the trends of those with and without obesity. It will inform the user and help gather additional information on the subject. The applications use of Planning, Tracking, Rewarding and Reporting keeps the user engaged throughout the process and aides in the collection of information for the research study.

1. Describe the process you will use to link the data from the different sources you’ve identified. Include a description of feasibility and any considerations that will be made to ensure the privacy, security and confidentiality of the data and data subjects throughout this process. (750-word limit)

Part of Research Kit includes access to Sage Biometrics Bridge Server - Sage’s Bridge Server is designed to securely manage data captured from IRB-approved human health research studies conducted through mobile technology platforms (iPhones and Android phones).

Bridge is a technology platform designed to support biomedical research studies conducted though smartphones and other sensor devices.

Bridge services support mobile registration and consent to participate in research studies, the design and scheduling of surveys and mobile sensor-driven activities, and the receipt of sensor and survey data from mobile sources.

Bridge SDKs are designed to integrate with the open source Research Kit (iPhone) and Research Stack (Android) frameworks to allow rapid and flexible development of research apps.

Bridge is developed and maintained by [Sage Bionetworks](http://sagebase.org/mobile-health/overview/), a not-for profit biomedical research institute active in the design, execution, and analysis of multiple mobile health studies. For technical support on Bridge or questions about this documentation please contact us at [bridge.info@sagebase.org](mailto:bridge.info@sagebase.org).

1. Describe how the linked data set(s) or individual data source(s) will be used to develop values for your proposed set of metrics in sleep, sedentary behaviors, nutrition, and/or physical activity. (500-word limit)

This method allows for researchers to examine trends acrous geographical, seasonal, age, or average phyiscal activity of users to aggregate data. This complex set of data is easily stratifiable and will help to improve research into obesity and heart issues as well as with the other studies areas. The data set will be hosted through sage bio networks.

1. Describe the representativeness of your data set for public health surveillance (e.g., to what population groups or sub-groups can you meaningfully extrapolate the results of your data set?). How amenable will this data set be to disaggregation by age, gender, education, geography, or other demographic characteristics? (750-word limit)

Part of the Research Kit onboarding includes the capture of demographic data of the subjects. The ubiquousness of the iPhone enables large scale research participation and as such provides a wide cross section of the population which can be broken down by age, gender, education, geography, or other demographic characteristics

We plan to ask as much information about the user as they are willing to answer to gain as precise a data set as possible. The app on first launch will ask a handful of survey questions to ascertain their demographic information. The data will allow for anonymity and for participants to choose to hide this information but will be marked as such so as not taint proper aggregated demographics.

1. How useful will your data set be for public health surveillance, how significant/relevant and generalizable are the results that you expect to obtain? (500-word limit)

The data that we will submit has very simple and transparent access requirements. This data can be used to make effective pushes on healthy activity to make an impact on the obesity epidemic. I nice side feature that our app adds to public health surveillance is the tracking of irregular heart beats. All these statistics will generalizable to allow for proper aggregation in a noSQL format. This format can be accessed by researchers, developer, and statisticians.

1. Statisticians can look at the data and crop out outliers and those that appeared to add chaos to the data set.
2. Developers can integrate the source into existing health surveillance networks and add onto the kit to expand it for more accurate research. Also they can integrate it into similar fields to acquire more data.
3. Researchers can gather the data and crunch the numbers in large computers or view them individually to judge as a group or demographic.
4. Will the proposed project’s data and data sets contain information of relevance to other areas of public health surveillance (e.g., chronic or infectious disease)? If yes, please specify and describe any additional work that would be required in order to expand applicability. (500-word limit)

Two major areas we are looking to build our app for is irregular heartbeats, and Obesity.

For obesity the major idea behind this is to gather accurate data on the user retention of those seeking to lose weight. We seek to gather information on both healthy and those that are unhealthy that may seek to lose weight. We believe that this is a major part of the obesity epidemic and that accurate information on the causes of drop rate in physical workouts offer vital information to curb it. On the other side information on the pickup rate of exercise and commitment to it can offer some insights on how to encourage those with obesity.

A side benefit of studying physical exercise is that we monitor people’s heartbeats. This can allow for us to detect for an irregular heart beat if we can aquire the proper algrothim. Which has already been developed.

1. Please describe a 3.5-month plan to develop a working prototype during the second phase of this challenge. This should include:
2. Details on how you will gain access to and link data from the source(s) you’ve identified.
3. Approaches/strategies that will be taken to ensure privacy/confidentiality of data before and after linkage.
4. Your approach to comparing results from your prototype to that generated from existing public health surveillance programs
5. A description of the format your prototype will take (e.g., visualization, online data tool, etc.)
6. Costs you expect to incur during this prototyping phase

(1500-word limit)

Our working prototype will come in the form of the development of the Healthy Behavior Apple Research Kit application. This application will be developed using standard agile software development practices.

Write about agile design where each sprint consists of design, code and test and produces a work product

We will use the agile development approach to software consisting of 2 week sprint cycles over the 3.5 months of development as follows:

Sprint 1: Wireframing / Design of the app

The first sprint will revolve around, integration of data sources and basic api features. We want to get as much of the back end code implemented in this phase. Some additional work will need to be included to match some of the screens and adjust the data. But the primary data sources will be set in this sprint.

Sprint 2: Development – Onboarding Screens of the application

On the Second sprint we seek to tackle UI aspects of the application. Research Kit conatins a lot of options for data collection and some not all of them will be nessecary for our application. Some additional refinements to the data storage solutions will be dealt with at this stage. Mock ups of screens will be implemented and sourcing of images will be taken care of here.

Sprint 3: Survey Questionaires Screen Development

At this stage in the process we seek to finalize screen’s for the majority of data gathering. The survey questions will play a key part dynamic data collection that traditional in person or phone surevys rely on. This should allow for deep analysis of stratified subject mentality on safety of working out and other hard to quantify data aspects. Short fo the health kit aspects the majority of the data collection should be taken care of.

Sprint 4: Health Kit integration Initial

This sprint will iniate the final aspects of active data recording. It will add in the initial stages of the final data gathering tool. We plan to add tracking features for heart rate to supplement survey questions and automated data gathering. We slowly prep the data for later integration into the bridge network sdk.

Sprint 5: Health Kit – Watch Kit Integration Final

Final modifications will be made on the Health kit, to fully implement the exercise data gathering set. This will revolve around capturing as much data as possible from the one touch exercise button. Some survey questions important to data gathering and garunteeing that user exercise is actively recorded on each occasion. Spuraditic recording of the data would messup the data slightly, so we compensate by trying to determine the active recording times to non recording sessions.

Sprint 6: Integration with Bridge Network SDK

This last few of sprints will be the final backend setup for the application. This will deal with issues keys and getting proper data storage methods up and running. This will also keep data secured with its high encryption. Note: While we will integrate with the Bridge Network SDK we will not implement the hosting solution or contract with Sage for them to host the data.

Sprint 7: Final Testing and Release

Privacy will be achieved through the integration of the Sage Bionetworks Bridge serveral although that layer of integration will not occur in the prototype phase.

The format of our prototype will be a Apple Research Kit app.

We expect to incur software development costs of $30,000

1. Significance and Relevance Summary

|  |
| --- |
| In 200 words or less, provide a brief summary of your project using language that is easily understood by the general public. Note: this description will be shared with a broad audience and should not include any information you would not want shared widely.  Healthy Behavior Research App will seek to improve the data tracking capabilities of the CDC. And supplement traditional sources with a sought after replacement to standard mundane data collection tools. To do this we plan to reward those that partipate with little badges of encouragement. We hope that our application can lead to better understanding to the causes of some chronic illnesses. |

Public reporting burden of this collection of information is estimated to average 60 minutes per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. An agency may not conduct or sponsor, and a person is not required to respond to a collection of information unless it displays a currently valid OMB control number. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to CDC/ATSDR Reports Clearance Officer; 1600 Clifton Road NE, MS D-74, Atlanta, Georgia 30333; ATTN: PRA (0990-0390).

1. Moderate-to-vigorous physical activity (MVPA) is any activity with an energy expenditure >3 metabolic equivalents [↑](#footnote-ref-0)