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. Details on the HBD Challenge may be found at challenge.gov.

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

Submission Deadline:

The Healthy Behavior Data Challenge
Phase 1 Submission Template

1) Challenge Team Information

Team Name
Onlife Health

Team Lead	City/Province
Catherine Bass, Ph.D.	Brentwood, TN
E-mail	**Phone Number**
cbass@onlifehealth.com	615-844-2186

Subject-matter/domain expertise
Measurement, program planning, needs assessment, health behavior theory, survey creation, public health, worksite health promotion

Team Member #1	**E-mail**	**Subject-matter/domain expertise**
Erica Zuhr, Ph.D.	**erica_zuhr@onlifehealth.com**	**Data science, predictive analytics**
Team Member #2	**E-mail**	**Subject-matter/domain expertise**
Tommy Strickler	**thomas_strickler@onlifehealth.com**	**Statistical modeling, data analysis**
Team Member #3	**E-mail**	**Subject-matter/domain expertise**
Jonathan Conley	**jonathan_conley@onlifehealth.com**	**Product development, mobile application design**
Team Member #4	**E-mail**	**Subject-matter/domain expertise**
Erik Hiser	**erik_hiser@onlifehealth.com**	**Data architecture, data modeling, database security**
Team Member #5	**E-mail**	**Subject-matter/domain expertise**
Sean McCoy	**sean_mccoy@onlifehealth.com**	**Mobile application development**
Team Member #6	**E-mail**	**Subject-matter/domain expertise**
Biaunca Edwards	**biaunca_edwards@onlifehealth.com**	**UI/UX, Graphic Design**

Are all team members residents of the United States?

Yes

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

Organization Name	Website	Type of Organization
Onlife Health, Inc (a.k.a. Onlife)	www.onlifehealth.com	Comprehensive Wellness Solutions

3) How did you find out about this challenge?

CDC website

4) Submission Overview

Project Title

Closing the Loop: Augmenting Mobile Data Sources for Public Health Surveillance

Project Overview

Describe in 500 words or less:

- What aspects of sleep, physical activity, nutrition, 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?

We propose to collect data points from all four reporting categories through a combination of device and mobile survey data. Data sources leveraged will include mobile phone sensor data (e.g., GPS, accelerometer), data from wearable fitness devices, and responses to cell phone administrated survey questions.

The wealth of data generated by mobile and wearable fitness devices is ideal for public health surveillance as it does not suffer from recall or self-reporting bias. Despite the aforementioned strengths, in our extensive experience mobile device data as applied to health behaviors is also incomplete, inconsistent and lacks context. We propose a series of short, just-in-time survey questions delivered via push notification to supplement mobile data sources.

Our combined methodology will improve on current public health surveillance in four primary ways.

1. Improve costs by reducing the human resource burden of administering the surveys
2. Improve recall bias by:
    - Leveraging wearable devices and mobile phone apps for objective data collection

- Administering just-in-time surveys to validate and provide context to device data and to collect additional data not tracked by wearable devices and mobile applications

3. Improve under/over reporting by leveraging wearable devices and mobile phone application for objective data collection
4. Improve response rate and reducing the respondent burden by:
    - making the questions more user friendly
    - creating engaging and easy methods of completing the survey

In order to access wearable device data, we will utilize Validic to de-identify and aggregate data on physical activity. Our organization already supports integration with the Validic API, and we have over 75,000 participants opted in to share their wearable device data. In addition to employing our exisiting data pipelines, we will build a mobile app to collect cell phone location data that will inform the timing and content of the survey questions. As an example, suppose a participant is identified as being at a fast food restaurant via mobile device data and cell GPS data. The user would get a text asking if they ate a meal there and a handful of follow up questions might include the number of servings of fruits and vegetables, number of sugary drinks, and number of servings of fried foods consumed. These questions take only a few minutes to answer, and provide context around the basic fact that the user was at a given location. As another example, if a long period of sedentary behavior is detected from a wearable device, the participant might receive a question asking if they were wearing their device during the sedentary period. If they respond in the negative, a follow up question regarding their activity level would then be delivered.

Following data collection, statistical analyses will cross validate the mobile and survey data sets with each other, as well as, the BRFSS survey data at the zip code or FIPS code level in order to understand consistency between all data sets. Depending on results, extrapolation of automated data collection at a much larger scale, without the additional expense of the survey methodology, is possible.

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

The data points we initially propose to collect are noted below.

A) Physical Activity

- Amount of MVPA[1] time per day
- Amount of MVPA time accrued while at work, at home and/or in transit
- Identification of times during the day where MVPA is high
- Frequency of MVPA
- Location of MVPA (recreation facility, at home, at work, on sidewalk/bike lane)
- Perception of safety while active
- Enjoyment level of MVPA
- Readiness to change[2] physical activity behaviors

B) Sedentary Behavior[3]

- Amount of time per day spent sedentary, excluding sleep time
- Amount of sedentary time accrued while at work, at home and/or in transit
- Number of hours spent in a car or motor-vehicle
- Readiness to change sedentary behaviors

C) Sleep

- Hours of sleep per night (sleep duration)
- Amount of time awake after sleep onset
- Consistency of bedtime
- Consistency of wake time
- Type of activity directly before sleeping (e.g., screen time)
- Sleep satisfaction in the morning
- Daytime sleepiness
- Readiness to change sleep behaviors

D) Nutrition

- How often fruit (not including juices) was consumed (day, week, or month)
- How many times per day/week/month a green leafy or lettuce salad, with or without other vegetables, was eaten
- How often vegetables (not including lettuce salads and potatoes) was eaten (day, week, or month)
- Number of sugar-sweetened beverages consumed in a week (or per day)
- Readiness to change nutritional behaviors

[1] Moderate-to-vigorous physical activity (MVPA) is any activity with an energy expenditure >3 metabolic equivalents

[2] Readiness to change is a measure assessed through the framework of the Transtheoretical Model for Stages of Change

[3] Sedentary behavior is any waking activity characterized by an energy expenditure ≤ 1.5 metabolic equivalents and a sitting or reclining posture

6) 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	Onlife/Validic	Wearable	API	Validic has fee for access which Onlife already subscribes to	Close to real time / collected multiple times per day	Physical activity, sedentary behavior, sleep, nutrition	Onlife, as well as, companies across the world. Validic aggregates data from more than 400 devices and has reached over 223M individuals
2	Onlife	Mobile app (accelerometer, geo sensors)	API	Open access APIs available	Close to real time / collected multiple times per day	Physical activity, sedentary behavior, sleep, nutrition	
3	Onlife	Just-in-Time Survey Responses	Specialized mobile application developed by Onlife	Variety of free and fee options to choose from	Close to real time / collected multiple times per day	Physical activity, sedentary behavior, sleep, nutrition	

7) **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)

Our solution will provide data that is both complementary to and significantly more novel than results from the survey methodology currently utilized by the CDC. All of our data will be indexed at the zip code, FIPS, and/or census tract level which will allow for comparison with the BRFSS data. Direct collection of a number of metrics that are tracked by the BRFSS (e.g., hours of sleep per night, servings of fruits and vegetables) and aggregation of our collected data into other BRFSS metrics (e.g., number of non-job related physical activity bouts per month) will make direct comparison with BRFSS results at these levels of geographic granularity straight forward. Our survey methodology will also allow for the collection of income, gender, and other demographic information that would be helpful as a data source complementary to CDC health surveillance data.

Our data source is more novel in two primary ways. The first of these is that we will utilize automated data collection through wearable and mobile data. Current CDC health surveillance does not utilize these rich sources of real time, unbiased data and Onlife has a great deal of experience pulling and aggregating this type of health behavior data. The second innovation of our framework is the use of the just-in-time survey messaging. Through this extensive experience with wearable and other mobile health data we can confidently say that the data received from the average wearable device has a large amount of missing data and, more importantly, has no context around the data collected. As an example, automated device data can tell us that a user is likely not wearing their device, but without context there is no way to know if an individual has taken off their wearable device for a one hour nap or a one hour lap swim. GPS sensors in a cell phone can tell a participant was at a McDonald's around lunchtime, but without context there is no way to know if the participant ordered a supersize meal with fries or a salad with no dressing. Our "closing the loop" methodology will not only collect wearable and cellular device data, but monitor it for missing or ambiguous data, push short simple questions to the participant based on this, and record the responses.

Lastly, this method of data collection does not require a human resource to gather data. This removes several limitations from the traditional data collection method from the cost of administering the survey (travel, surveyor salary) to the limitations on the number of surveys and questions that are able to be collected. Rather, collecting data in a more continuous manner, through the use of devices, enables us to collect more data, more objective data and may provide a more comprehensive view of the member's behaviors than a one-time response.

8) **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)

Onlife is HIPAA compliant. We currently hold two URAC designations and one NCQA accreditation, both requiring HIPAA compliant policies for all PHI or potential PHI data. Additionally, we are a Health Information Trust Alliance (HITRUST) CSF Certified organization. Holding this certification shows Onlife has a framework in place that has the needed structure, detail and clarity in relation to information security tailored to the healthcare industry.

Onlife takes appropriate security measures in regard to equipment, software, intrusion control, virus detection, etc., to ensure data security. Appropriate hardware and security configurations, as well as role-based active directory, specialized enrollee management user groups and encryption, prevent unauthorized access. Those with access to the participant management system are tracked by each area they access in the system and for which enrollee they are accessing it. All customer data accessed is logged, so that if there are any questions, they can be resolved from the log records.

The most secure data transmission and encryption standards are used to ensure the integrity of all data entrusted to Onlife. Internal security scans are conducted semi-annually along with an external audit every eighteen to twenty-four months by a third party with expertise in that area.

Internally, data will all be stored at the participant level, and wearable data will be linked using a unique person ID to cellular sensor data and mini-survey responses.

For any reporting or retransmission for further evaluation and aggregation by the CDC, all identifiable, non-required fields will be non-included. This de-identification will occur according to general HIPAA guidelines, to appropriate minimum-to-show levels for a given geographic location, or other demographic evaluator as currently dictated by CMS guidelines, except where CDC needs or requirements differ.

9) **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)

A number of the metrics we propose to collect Onlife already pulls and aggregates through our partner, Validic. As an example, hours of sleep per night is available for almost all wearable devices through Validic. For participants who do not have a wearable device, we will encourage collection of this data point through health data aggregation apps already integrated with Validic such as HealthKit and GoogleFit. Finally, on nights when data is not collected through any of these sources (or our data quality indicators flag data collection as poor or incomplete) the participant will be pushed a short series of questions. A potential question flow appears below.

- If participant registered sleep for previous day
    - If (yes): "Does (x) hours of sleep seem accurate to you?"
    - If (no): Ask correct number of hours, with scroll bar or spin button input starting from the number we have registered.

- If participant did not register sleep for previous day
    - Were you wearing your device?
    - How many hours of sleep did you get last night? (Spin button input starting from user average).

Minutes of moderate to vigorous physical activity (MVPA) per day will be collected via wearable device through Validic in a similar manner. MVPA is available for Fitbit and Garmin devices, which make up the vast majority of consumer wearables used today. We will supplement this data with a handful of short survey questions whose number and content will depend on the quality of the data we receive from Validic. A sample question might be "We registered 37 minutes of MVPA for you at home yesterday. Is this correct?" If the data is correct the participant simply clicks yes to verify, and if not they will be prompted to adjust the time or location to create a more accurate data point. Finally, the user will be asked if they have any additional bouts of MVPA to enter.

Nutrition data is also collected through Validic, in the form of a food diary recorded through fitness apps or native operating system health aggregators. Most of the answers received in this field are free response, so our team will parse the collected for a pre-set collection of fruit, vegetable, and sugary drink responses using standard natural language processing techniques to cast the widest net possible. We will push any data to the user to verify, give them the option of entering additional data points, and if no data is collected request that information via simple multiple choice questions.

Data collected in this manner can be completed in a matter of seconds with just the swipe of a finger, which reduces reporting burden on the public and increases response rate. The comparison and basis in unbiased device data should reduce recall bias and over/under reporting. Learning a user's history and having autocomplete available will also allow us to push more intelligent questions that reduce participant burden in this area.

10) **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)

In general, we expect our data set to be largely representative of the adult population in the United States. Onlife's membership has over 75,000 participants who have connected their devices to our data. Our members range in age from 18-101, with an average age of 42. Males and females are represented fairly equally, with females being more highly represented at 61%. Our members are representative of all 50 states, with the densest populations being in the southeast, northeast and midwest. Members have a wide range of education levels, from those who have less than a high school diploma to those with graduate degrees. The average education level among the membership are those who have completed some portion of college. Various races are represented in our membership with non-Hispanic Caucasians representing the largest majority, followed by African Americans, Hispanics, and Asian/Pacific Islanders. These demographic characteristics are reasonably consistent with the population at large so that generalizations can be made with confidence.

We do have one bias in our population that should be noted. That is that Onlife members are adults who have health insurance coverage – although they may or may not be the primary beneficiary. This particular bias in sampling is not inherent to our framework, and in future iterations with CDC involvement, will be easy to eliminate.

As Onlife serves members from every state and across a wide variety of age groups and demographics, we can work with the CDC to select a sample for the prototype that is comparable to the sample from the BRFSS. We have a large enough population that our dataset can easily be disaggregated by age, gender, education and other demographic characteristics.

For our initial proposal, we do intend to track only adults with smartphone access, but research into prevalence of mobile technology shows that this will not be a barrier to collecting a representative sample for an initial prototype or a significantly larger study. The vast majority of the US adult population already has access to such devices, with the Pew Research Center reporting 95% of adults owning a cell phone, while 77% own a smart phone. In addition, Pricewaterhouse Coopers reports 1/5 of Americans own a wearable health device.

The smartphone limitation is not inherent to the framework of the proposed methodology and need not limit the CDC's methods going forward. Although we will not incorporate this feature in our pilot, adding participants with a cell phone that is not a smart phone is feasible through straight forward adjustments such as delivering survey questions via text message instead of mobile app. This adjustment could help increase the representativeness of the sample by using data from cell phone sensors and health data aggregators. While this data is less specifically tailored towards health behaviors, it can be supplemented with additional SMS survey questions to collect any missing data.

We do plan to include participants who own smartphones but not wearable devices in our pilot, collecting data from apps and using push notifications to administer survey questions. Data quality for some items for those participants may be less robust and more subject to the biases of traditional survey methodology. While this may initially be more limiting as we will request users to "bring their own device" for participation in our prototype, in the future the number of users with devices in the general

population will likely increase, and as the price of wearables continues to decrease it may be feasible to purchase devices for participants as an incentive.

11) **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)

With 77% of the US adult population owning a smart phone[4], reaching sample sizes on survey questions to achieve acceptable statistical confidence and margin of error across demographics (age, gender, geography, etc.) is reasonable by employing a marketing campaign for the application, using proven recruiting techniques, such as offering an incentive, and leveraging partnerships with local and state health departments.

Generalizing survey results and metrics retrieved via wearable devices to the US adult population can be accomplished using statistical weighting techniques (e.g., age-adjusted prevalence rates), similar to current BRFSS data analysis methodology.

As an additional important feature, we believe the results of this initial study will be generalizable to other mobile data collection studies that may not have the time or resources to supplement device data with survey questions. In addition to statistically significant results on the population we are following, we may be able to better understand what device data means at a larger scale – for example, if a large population has an average of 35 minutes of recorded MVPA a day, perhaps we can show that, in actuality, that population may be getting closer to 60 minutes of activity a day.

Depending on the statistical significance of the data collected, these types of generalizations may be applied at the gender, age or other demographic level. Perhaps women tend to record more of their actual activity than men, or older people wear their devices more consistently during the day but less consistently at night. These insights will be invaluable for a field just beginning to utilize the wealth of incomplete, inconsistent health data being collected at scale by many major companies and devices.

[4] Pew Research Center, http://www.pewresearch.org/fact-tank/2017/01/12/evolution-of-technology/

12) **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)

Our proposed data collection framework has elements that could be useful to other areas of public surveillance, and could easily be extended to be even more specific to other areas of interest. A literature review shows that much of the existing research on SMS or cell phone surveys is based in epidemiology. Response rates were promising, even without additional financial incentive for study or survey participants.

This methodology is immediately relatable to chronic conditions, as many chronic conditions benefit from improved lifestyle behaviors. Understsanding more about the populations' behaviors can inform campaigns designed to improve chronic conditions. This information also lends itself to monitoring progress in improving chronic conditions. Additionally, survey questions can be easily modified to fit the subject matter for collection and validation of data.

Thinking in terms of infectious diseases, pushing a simple survey question through our push notification to all users around epidemics would be one useful application, but we could also extend our monitoring to push questions based on location. If a participant is near a doctor's office or hospital, the app can push a quick question to verify their location, and then ask if their presence at the medical building is related to a particular disease or set of chronic conditions that the CDC wishes to track. We believe that the general framework of passive data collection combined with a limited number of active questions, to close the loop, has nearly unlimited potential.

13) Please describe a 3.5-month plan to develop a working prototype during the second phase of this challenge. This should include:

1) **Details on how you will gain access to and link data from the source(s) you've identified.**
2) **Approaches/strategies that will be taken to ensure privacy/confidentiality of data before and after linkage.**
3) **Your approach to comparing results from your prototype to that generated from existing public health surveillance programs**
4) **A description of the format your prototype will take (e.g., visualization, online data tool, etc.)**
5) **Costs you expect to incur during this prototyping phase**

(1500-word limit)

Our final deliverable will be a cleaned, aggregated and easily accessible dataset merging wearable device data, consumer health mobile application data, and the responses to survey questions which are closely related, and in some cases, dependent on the wearable and mobile data points. We will also produce statistical analyses of this data both as an independent dataset, and utilizing direct comparisons to BRFSS datasets. Linkage of all datasets will be at the user level, and can be aggregated by census tract, FIPS code, zip code, or any other larger geographical designation to compare to BRFSS data. Most of our data points either directly mirror or roll up into a question that is asked on the BRFSS, which will allow for simple correlation analysis of results.

All mobile and wearable data will be collected through Validic, a data integration service with whom Onlife has an existing contract and existing data warehouses. Validic is compatible with hundreds of wearable health trackers and devices, and Onlife currently pulls the majority of their data points we wish to collect for this prototype. Our system currently contains wearable device data on hours slept at night, minutes of MVPA and sedentary minutes per day as well as user entered nutrition data. We will continue to access wearable data as our system has for years.

Included in this final porotype will be a mobile application that serves both Android and iOS platforms. The app will monitor data in the background and on our servers, using this information to push and collect the answers to just-in-time survey questions. Data confidentiality will be ensured by adherence to internal data security policies and procedures as described in the previous section, as well as adherence to national agency guidelines surrounding identifiable data and related minimum to show guidelines for geographic and/or demographic features.

Weeks 1-4

- Select client[s] and participants to target for prototype
- App development
    - Login Screen UI
    - New User Screen UI
    - Brief Questionnaire to classify user Screen UI
    - Survey Questions Screen(when getting questions) UI
- Select survey questions and finalize wording
- Finalize data model for survey responses
- Review data model for Validic data, adjust if needed and create plan for joining with survey and other data

- Statistical review of current Validic data to help create data quality metrics that will guide survey question delivery

Weeks 5-6

- App development
    - Back end services
    - Service layer in the mobile item
    - Database connection and Stored procedure creation
    - Debugging and QA
- Finalize criteria/flow charts for push notification survey questions
- Begin participant recruitment

Weeks 7-8

- Continue to recruit and opt in participants
- Final UAT for mobile app, push to app store

Weeks 9-12

- Collect data through devices and just-in-time survey questions
- Create analysis pipelines for incoming data

Weeks 13-14

- Clean and analyze the initial month of health data collected through devices and survey questions
- Aggregate daily values into metrics that compare with BRFSS questions, typically at a monthly level. Include statistical metrics for confidence and significance
- Baseline comparison to BRFSS data
- Initial disaggregation by age, gender, and other demographics factors

Costs:
Onlife employee time will be the primary cost for our prototype, and we have committed several resources from our informatics, data services, software development and product teams who will have time dedicated to this project during the 3.5 month prototype development period. Onlife will also offer a small incentive to participants as part of our recruitment for the study. We will aim to recruit 350 participants to account for attrition. Participants will be given a $5 gift card in exchange for opting in to participate and another $10 gift card if they continue participation throughout the study period. Assuming maximum participation, there will be a cost of $5,250 dollars in incentives.

While there will be no cost for data aggregation through Validic during the prototype period as we will use our existing contract, we strongly suggest a continued usage of Validic's services in data aggregation and deidentification which may be an additional cost to the CDC in the future. It should be noted, however, that custom API connectors and opt in technology can be developed for each app the CDC desires to use in future studies, and if the number of apps or devices utilized is limited this may be a cost effective alternative to Validic.

14) 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.

Wearable devices and mobile applications are an invaluable source of data for public health professionals to tap into in order to gather and monitor the behaviors of the public at large. Gathering data from this technology requires relatively little effort for both the public health professional and the individual by making use of technology the individual has already chosen to use.

While data gathered through devices and apps is often objective and reduces error around recall bias and under/over reporting, it can also be incomplete and void of context. For example, when trying to understand an individual's physical activity habits, it may appear that they have not had any MVPA minutes in a day. Using this data to push a question about their physical activity may help validate their lack of MVPA or may bring to light that their chosen exercise is swimming, which their device does not track. Interweaving qualitative data with the data gathered from devices and apps provides context and validation. Couching data in this way allows for a fuller understanding of an individual's behavior, which is critical for public health officials who are trying to monitor the nation's health, identify needs and make decisions around initiatives.