Study Protocol and Statistical Analysis Plan

Using Retrospective and Real-Time Physical Activity Tracking to Predict Risk of Sunburn in Outdoor Exercisers on Strava

NCT04977700

06/22/2020

Using Retrospective and Real-Time Physical Activity Tracking to Predict Risk of Sunburn in Outdoor Exercisers on Strava

1R21CA241637-01A1

Recreational UV exposure is associated with every form of skin cancer and individuals who engage in more physical activity have a higher prevalence of sunburn, a proximal biomarker of melanoma risk, perhaps explaining why melanoma is the only cancer with which physical activity is positively correlated. Mobile technology for tracking physical activity has become increasingly prevalent and Strava, an activity tracking app and social networking site for athletes, is one of the most popular of these technologies. This research will test the feasibility of delivering location-based, ecologically-valid sun safety advice to Strava users at times when they are predicted to be engaged in outdoor physical activity, by utilizing Strava's public open-source Applications Programming Interface.

2. SPECIFIC AIMS

Skin cancer is the most common cancer in the United States. *In 2019, over 5.4 million cases of keratinocyte cancers (basal and squamous cell) and 96,480 cases of melanoma will diagnosed and melanoma will cause 7,230 deaths.*[3] Skin cancer treatment is a financial burden on the nation's health system.[3] Thus, the U.S. Surgeon General has designated skin cancer prevention as a major national priority.[4] Skin cancer is preventable because excessive exposure to ultraviolet radiation (UV), the primary risk factor, is easily modifiable. But, many Americans do not apply sunscreen, wear UV-protective clothing or wide-brimmed hats, or stay in shade and approximately one-third of American adults experience damaging sunburns annually.[5]

Recreational UV exposure is associated with skin cancer.[6] Individuals who engage in more physical activity have a higher prevalence of sunburn,[5,7-10] a proximal biomarker of melanoma risk, and melanoma is positively associated with physical activity.[11] For some athletes, forgetting to apply sunscreen[12,13] or not liking its feel while engaged in physical activity can be barriers to sun protection.[12,14] An expert meeting hosted by the National Cancer Institute (NCI) identified sun safety during physical activity as a priority research area[15] and the Surgeon General prioritized strategies for coordinating messages on sun safety and physical activity in his *Call to Action to Prevent Skin Cancer*.[16]

Online and mobile technologies for tracking physical activity have exploded over the past decade. Many Americans wear activity trackers or GPS-enabled devices (e.g., Fitbit and Garmin) and upload their physical activity data to websites/mobile apps. *Strava*, an activity tracking app and social networking site for athletes of all abilities, is one of the most popular. Millions share their own activities and follow and comment on physical activity by other users in their network. *This R21 research responds to PAR-19-309 "Stimulating Innovations in Behavioral Intervention Research for Cancer Prevention and Control" with the* ***goal*** *of increasing skin cancer prevention among a high-risk population, adults who engage in outdoor physical activity* by establishing **feasibility** of interfacing sun protection advice with the *Strava* website/mobile app. Specifically, we will show that we can create an algorithm that a) predicts when individuals are likely to be engaged in physical activity outdoors and b) delivers sun safety advice tailored to time, location, and personal risk (e.g., skin sun sensitivity). We need to establish feasibility before conducting a future randomized trial evaluating the efficacy of this unique intervention. The *Strava Sun* (*SS)* program will obtain user data and deliver ecologically-valid sun safety advice by utilizing *Strava*'s open-source Applications Programming Interface (API) and location-based advice algorithms developed earlier for a sun safety mobile app, *sunZapp*, funded by the NCI.[17] Sun safety messages provided through *SS* will also be derived from sun safety messages from our *Go Sun Smart (GSS)* program for individuals engaged in outdoor recreation.[18] Both the algorithms from *sunZapp* and the *GSS* messages improved sun safety practices of adults in general and while on vacation in randomized trials (see **Boxes 1 and 2**) and provide the **scientific premise** for developing *SS*. *Based on the Extended Parallel Processing Model*[19] *of persuasive communication, Diffusion of Innovation Theory*[20] *(DIT), and Social Cognitive Theory (SCT),*[21] *we* ***hypothesize*** *that SS will increase sun protection by elevating perceived risk, self-efficacy, and response efficacy and promoting prevention skills acquisition.*[21] The specific aims are to:

1. Gather formative data from focus groups of outdoor exercisers, both *Strava* users and non-users, on interest in receiving personalized, location-based sun protection advice and willingness to use sun protection advice sent through the *Strava* mobile app before and during physical activity.
2. Develop a machine learning algorithm to predict future time of day and duration of outdoor physical activity from activities logged on *Strava* through its API from users who consent to data access.
3. Modify the algorithms in *sunZapp* and social media messages from *GSS* to create a prototype of the *SS* intervention that delivers location-based, tailored sun safety advice to *Strava* users through *Strava* comments, which can be accessed through the *Strava* website and mobile app, and emails.
4. Test with *Strava* users the acceptability of the *SS* prototype and sun safety messages, ways of interacting

with it, and barriers/willingness to allowing access.

*The R21 research is **significant** and **innovative**. A sun protection interface for the Strava platform will allow us to identify a large population of Americans at high risk for skin cancer who routinely engage in physical activity, often outdoors with high-risk sun exposure, and reach them with sun safety advice they may not seek on their own to motivate them to practice sun safety during outdoor activities. Dissemination of evidence-based sun safety messages through this widely-used "social network for athletes" is innovative. It will help to meet the goals set by both NCI and the U.S. Surgeon General to prevent skin cancer among physically-active Americans to reverse the one cancer currently increased by physical activity.*

The project will be led by Principal Investigators, Dr. David Buller and Ms. Julia Berteletti at Klein Buendel, Inc. (KB). *Co-Investigators at Colorado State U., Dr. Kimberly Henry (project statistician) and Dr. Chuck Anderson (machine learning expert) will create the predictive algorithms, assisted by a machine learning graduate student.* Dr. Sherry Pagoto, *project consultant*, at U. of Connecticut will provide expertise on physical activity promotion and interventions delivered via social media. Mr. Scott Camichael, recent software engineer at *Strava*, *will also serve as a project consultant* and advise on delivering the sun protection intervention through *Strava*. Media developers at KB will produce the API database and program.

3A. SIGNIFICANCE

3A.1. The Skin Cancer Problem

The research is significant because it will develop a new mHealth strategy to reduce skin cancer. In 2019, 5.4 million cases of keratinocyte cancers (basal and squamous cell)[4] and 96,480 cases of melanoma will occur. Melanoma has doubled over the last 3 decades[22] and will cause 7,230 deaths this year[23] (nearly as many deaths as due to drunk driving[24]). Exposure to solar ultraviolet radiation (UV) is a primary cause,[25-34] both lifetime UV exposure[25-28,35] and intermittent, severe exposure (i.e., sunburn).[15-21] The U.S. Surgeon General issued a *Call to Action to Prevent Skin Cancer*,[16] citing its high prevalence[4] and annual cost ($4.8 bil. for non-melanoma skin cancer; $3.3 bil. for melanoma[36]). Its recurrence,[37-39] and treatment disfigurement[40-43] also prioritizes prevention. Primary prevention of skin cancer relies on reducing UV exposure by limiting time in the sun when UV is high (e.g., at midday; at lower latitudes; near summer solstice) and using shade, protective clothing, and sunscreen. But, people inconsistently take precautions,[16] so efforts are needed to promote sun protection.[44-46]

3A.2. Correlation between Physical Activity and Skin Cancer

The research is significant because it seeks to improve sun protection among individuals engaged in physical activity, a beneficial health behavior that is also related to unprotected sun exposure and risk for skin cancer. Meeting guidelines for physical activity provides numerous positive health effects;[47-51] however, melanoma has a positive risk association with physical activity.[11] Physical activity is associated with sunburn,[7,8,52-54] skin cancer risk behaviors,[9] melanoma in both genders,[11,55] and keratinocyte skin cancers, chiefly in men.[56,57] It was estimated that sunburn is twice as likely to occur in physically-active than inactive people[52,53] and is linked to aerobic activity (e.g., running and cycling).[5,53] Cardiorespiratory fitness is related to an increased risk of melanoma[55] especially in high-UV locations, suggesting sun exposure plays a role.[11] In one study, persons who were more physically active were less concerned about sun exposure.[58] Together, the evidence implies that the behavioral phenotype of exercisers leads to increased UV exposure, elevating risk for developing skin cancer.

Promoting sun safety to exercisers must not discourage physical activity. It should be ecologically-valid by considering the person's age, environment, activity type, and other sociocultural factors.[17,59] For example, we should promote sun protection as a personal comfort and safety issue; bundle it with factors like injuries (sunburn is an injury), heat stress, and hydration; note sun safety does not interfere with and may improve performance; and promote waterproof sunscreen and hats to avoid sweat leeching sunscreen into eyes.

3A.3. *Strava* as a Platform for Skin Cancer Prevention

Adding to significance are plans to use *Strava*, a popular social networking and activity-tracking site for athletes of all abilities, to reach high-risk individuals engaged in outdoor physical activity with sun protection advice tailored to time, location, season, and personal risk. *Strava* is a free-to-use networking website with mobile apps (Android and iOS), currently used by *42 million athletes, with approximately 1 million new users every month.*[60] Most activities are tracked using *Strava's* GPS features, *through a GPS-enabled watch (e.g., Garmin; Apple Watch) or Strava app*. Analytics data provided by Mr. Carmichael, project consultant, show that 32% of *Strava* users identify as runners, 32% as cyclists, and 22% as multisport (i.e., triathlon). Per US Bureau of Labor Statistics, 9% of American adults are runners and 3%, cyclists.[61] Running, cycling, and triathlon *(including swimming)* require large amounts of time in the sun, often exceeding sunburn doses of UV (i.e., ≥1 minimal erythemal dose [MED]). *Currently, ~90% of activities recorded are performed outdoors (~15 million per week[62]), using real-time GPS. Uploading an activity to Strava can be done nearly automatically, by simply choosing activity (e.g., running, cycling), starting a GPS-watch or Strava app, clicking 'done' when finished. The activity syncs to Strava servers using a small amount of cellular data. Users can add more detail to the activity, such as title, description, and photos. Strava* is also an online community for athletes. Using a feed similar to *Facebook*, activities of followers are highlighted and users can give "kudos" and/or leave comments on activities. This social nature encourages members to use *Strava* regularly and not just for personal tracking. *Strava users are accustomed to receiving comments[63]; in one study, 80% of Strava users received social feedback on their activities.[64] Receiving social support in an online fitness community increased levels of use.[64] In 2017, Strava users left 110 million comments and gave 2.3 billion kudos.[65]*

The *Strava* open-source Applications Programming Interface (API) is a publicly-available interface that is open to registered *Strava* users (at labs.Strava.com/developers) and allows developers access to *Strava* data.[66] A *Strava* API application only has access to a user's personal data if the user authorizes it. The API is an extremely promising and innovative tool for physical activity and sun protection. In this R21 project, we will determine if we can use the API to a) download retrospective activities and create an algorithm that predicts when users will be outdoors and b) deliver sun protection advice tailored on time, location, season, *and personal risk* to users through their *Strava* feed, accessable any time on any device. *Per Mr. Carmichael, the Strava API is constantly evolving and we are confident we can access the data needed through the API.*

There is no comprehensive theory that explains how mobile interventions improve health behavior.[67-69] As in our past projects,[70-76] Strava Sun (SS) will follow principles of the Extended Parallel Processing Model[19] of persuasive communication, Diffusion of Innovation Theory[20] (DIT), and Social Cognitive Theory (SCT):[19,21] a) communicating real-time (e.g., high UV hours, 10 AM to 4 PM), location-based advice to create perceived

threat to motivate prevention; b) teaching/promoting simple ecologically-valid sun safety skills with how-to knowledge to elevate self-efficacy (e.g., when to apply and reapply sunscreen); and c) showing why these practices prevent skin cancer (i.e., principles knowledge) to increase response efficacy. SS messages will remind users to take precautions when exercising outdoors and convert intentions into action.

3B. INNOVATION

3B.1. Innovation of the Proposed Research

The major innovation is a program that uses an open source API to predict outdoor physical activity bouts and provides location-based sun safety advice through the *Strava* tracking and social networking site and app.

- *Strava Sun* (*SS*) will utilize a popular platform to deliver evidence-based, temporally-based, location-based, ecologically-valid sun safety messages to individuals before they participate in outdoor exercise.
- We will use machine learning techniques to develop an algorithm using *Strava* activities to predict high-risk behavior, (i.e., outdoor physical activity when UV levels are high and sun safety is advised).
- *SS* will contribute to the mHealth and integrative data analysis research on utilizing publicly-available APIs and large consumer databases to predict high-risk behavior and access social media platforms to reach at-risk populations with technology-based tailored interventions.

This R21 project is high-risk, high-reward research. We are unsure that users will engage with and benefit from a location-based, real-time sun safety intervention delivered by the *Strava* service, but our *sunZapp* mobile app and *GSS* outdoor recreation program suggest they will (see **Boxes 2** and **3**). Over 80% of Americans own a smartphone and most use social media and mobile apps.[77-79] Findings from this R21 need to confirm the potential of *SS* before investing in a randomized trial. With the growing number online activity trackers (e.g., *Strava, Fitbit, MyFitnessPal*), it is difficult to compete for users' attention with a new stand-alone app for skin cancer prevention. *SS* will have high impact by using a popular established activity tracking platform to predict high-risk behavior and deliver evidence-based, tailored advice to a large population of outdoor exercisers.

Box 1: Preliminary Studies: Surveys and Focus Groups with Physically-Active Adults

Our team has collected formative data relevant to the proposal. Dr. Pagoto surveyed 278 adults who engaged in moderate-intensity physical activity. Physically-active participants had more sun exposure, higher perceived skin cancer risk, but similar sun protection as inactive adults. There were 4 types of active participants: 1) low perceived risk, no tanning interest, and high frequency of exercise in peak UV (17%); 2) low perceived risk, no tanning interest, and avoid exercise during peak UV (36%); 3) high perceived risk, no tanning interest, and avoid exercise during peak UV (18%); 4) high perceived risk, high tanning interest, and exercise during peak UV (29%). A second online survey assessed members of a large workout group (n=60; 91% non-Hispanic, 86% White, 67% female) in Denver in 2018. 94% exercised outdoors for ≥30 minutes at least 3 times per week (51% 5 times/week). 53% did so between 8:00am and 7:00pm and 16%, 10:00am and 4:00pm when UV is highest. 93% used a device to track physical activity, with a GPS-enabled watch most common (48%). 82% uploaded their activities on a tracking site (61% used *Strava*). Of *Strava* users, 78% regularly or always uploaded their activity; 43% uploaded 2-4 times per week, and 41% ≥5 times per week; 51% checked *Strava* notifications regularly. 73% were somewhat (51%) or very (22%) willing to get sun safety messages on *Strava*; 89% were somewhat (54%) or very (35%) willing to download an app that connects with *Strava* and uses their data to provide sun safety advice. Respondents, on average, received 2.6 sunburns in the past year (range: 0-20), with 2.5 of these sunburns occurring while exercising outdoors (range: 0-15).

We piloted our focus group protocol with Strava (n=5) and non-Strava (n=5) outdoor exercisers (90% non-Hispanic White; 70% female; mean age=32). Strava users exercised outdoors 4-10 times per week. They uploaded all of their activity to Strava, checked Strava daily for notifications, and were members of 4-6 clubs. They believed Strava protects their data. Most wore hats while exercising outdoors, but other sun protection was uncommon. All were interested in getting sun safety advice through Strava, especially UV Index, high-UV warnings similar to poor air quality notices, and cumulative sun exposure based on time spent exercising. They said the proposed SS messages would make them consider sun protection before exercising outdoors. Non-Strava users exercised outdoors 4-7 times per week. All applied sunscreen to their face but did not reapply it. Some wore hats but not UV-protective clothing or sunglasses. They did not use Strava because they did not want another social network and were not competitive (only 1 avoided it for privacy), but they would be interested in sun protection information from their weather app. Discomfort with sun protection was a barrier for both groups. Strava users also cited inconvenience and "laziness" and non-Strava users, not wanting to carry items and social norms as barriers. In focus groups conducted by Dr. Pagoto with adults (n=15; 52% female) who met physical activity guidelines (moderate-intensity activity >150 minutes per week), most common reasons for not using sunscreen (cited by 88%) were forgetting, inconvenience, discomfort, and belief they do not need it. Most common barriers to protective clothing were discomfort (e.g., too hot), cost, and desire to get a tan while exercising.

Scientific Premise: *This research verified that many outdoor exercisers are a high-risk population, regularly receiving dangerous doses of UV with inadequate sun protection. It also confirmed that activity tracking is popular and Strava users regularly and frequently log activities, check notifications, and are interested in receiving sun safety messages through Strava using SS. Our intervention messages will be tailored to address content needs and barriers to sun safety cited by these outdoor exercisers.*

3C. APPROACH

3C.1. Project Overview and Timeline

In the two-year R21 research (see timeline in **Section 2.7. Form E**), we will conduct rigorous formative research and pilot-testing of the *Strava Sun* (*SS*) intervention. We will create study protocols in Month 1. Outdoor exercisers (*Strava* users and non-users) will give input on initial concepts for *SS* in focus groups during Months 2-5. An algorithm that predicts time of day and day of week exercising outdoors during high UV will be built using retrospective *Strava* user data in Months 6-8. In Months 9-15, a prototype of the *SS* database will be created and messages and delivery methods tested for usability/acceptability. In Months 16-21, feasibility of *SS* will be tested in a pilot field trial with *Strava* users. *While the many aims are to establish methodologic feasibility, analysis of the pilot field trial in Months 22-24 will test the* ***hypothesis*** *that SS will improve sun protection. To achieve the ambitious timeline, we will use protocols and methods developed in prior research (see* ***Box 2*** *and* ***3****).* All procedures will be approved by the Western Institutional Review Board.

3C.2. Target Population: Outdoor Exercisers on *Strava*

Target population will be users of the *Strava* social network. Inclusion criteria include a) being 18 or older, b) U.S. resident, c) engaging in regular physical activity outdoors (at least 156 times in past year [3 times per week]), and d) consenting. All participants except non-*Strava* users in focus groups will: e) be tenured *Strava* members (for 1+ full calendar year), f) upload at least 13 activites to *Strava* in past month (i.e., 3 per week), and g) give permission for us to access their activity data. *Strava* has *42 million members* worldwide, providing a very large user population. *All recruiting, besides for focus groups, will occur by invitations through Strava (e.g., posting in clubs). Interested users will click a link to a study registration website. We cannot screen users based on inclusion criteria on Strava directly, so interested trial participants will complete a screener on the study website to determine eligibility. We expect participants to regularly participate in outdoor exercise, thus being a high-risk group, and to actively use Strava.*

3C.3. Formative Focus Groups (Specific Aim 1)

Four **focus groups** will be conducted with adults who engage in outdoor physical activity and use *Strava,* or not, in Colorado in Months 4-5 (compensation=$50). *Using Facebook and investigators' networks,* we will recruit a variety of exercisers, including those who use *Strava* for recreation only and for competitive training. Though not mutually exclusive, we will attempt to segment groups by these uses. *Strava* users (n=2 focus groups; n=16 total; 8 male, 8 female) will discuss a) *Strava* use, b) exercise behaviors, c) concerns about privacy, d) willingness to receive sun safety advice while exercising and *e) use of sun protection while exercising*. To explore interest in sun safety advice by those who track physical activity through other means, non-*Strava* users (n=2 focus groups; n=16 total; 8 male, 8 female; *matched to Strava users in focus groups on race, gender, and activity type*) will discuss exercise behavior, privacy concerns, and why they do not use *Strava*. Both groups will discuss time of day/day of week and season they exercise outdoors, activity type, sun safety intentions, tanning desirability, barriers to sun safety while physically active, *SS* concept, preferred sun safety content/delivery methods, barriers to using *SS*, utility/acceptability of sun safety advice through *Strava*, potential competitive advantages of sun safety, willingness to use the advice while exercising, and privacy concerns. We will not segment focus groups by age, but will examine age for differences in tailoring preferences. Transcripts will be coded and analyzed using ATLAS.ti.® *For this resubmission, we pilot-tested the focus group protocol (****Box 1****).*

3C.4. Development of Algorithms (Specific Aim 2)

Dr. Anderson, Dr. Henry, and a computer science graduate student will develop a model that can ***predict*** *impending outdoor physical activity by Strava users during high UV based on their uploaded activity records and relevant ambient factors (e.g., weather) and* ***prescribe*** *timely sun safety advice.* We will use data-driven machine learning analytics to develop an algorithm that forecasts the next outdoor physical activity bout, assuming past behaviors and situational contexts predict future behavior.

3C.4.1. Collect Retrospective Activity Data. Data from 1,000 tenured adult *Strava* users in the 4 U.S. Census regions will be obtained. *Eligibility criteria includes: 1+ full calendar year of activities; at least 156 activities [3 per week] in past year; 18 and older; U.S. resident.* Users will consent for their data to be accessed and a full calendar year of all *Strava* data will be procured, including following variables, *which are*

all available through the Strava API[80]: distance, moving time, pace, elapsed time, start time, end time, date, type of activity (e.g., run; ride; *swim*), and gear used. These data will be paired with calendar data (i.e., day of week, season, holidays) and weather data *from National Oceanic and Atmospheric Administration databases* (i.e., temperature; precipitation; humidity; UV level; wind) to create a multivariate time series of data for each user. This data will be subjected to machine learning analytics to develop an algorithm that forecasts the likelihood that the user will engage in outdoor activity *in near future (duration of reliable prediction will be modeled)*, including the nature of exercise bout (e.g., length of time; UV level). We will recruit users by posting in clubs, a popular *Strava* feature (e.g., *Strava* Club has over 662,800 members[81]). We will post in various clubs until 1,000 eligible members are enrolled. Participants will receive an authentication screen (see **Figure 1**) that identifies data we will access and how it will be used. We will raffle off 10 $50 gift cards as an incentive.

3C.4.2. Use Machine Learning (AI) to Predict Outdoor Activity. Deep learning analytics, a branch of machine learning that mimics structures of the human brain, are innovative techniques that give more flexibility and are typically more accurate than autoregressive integrated moving average (ARIMA) models.[82,83] To predict future outdoor activity, we will use a type of artificial neural network, Recurrent Neural Network (RNN), well-suited to temporal prediction,[84] along with a Long Short-Term Memory (LSTM) approach.[85] LSTM neural network can learn long-term patterns and dependencies and provide stability to prediction algorithm by evaluating when it is advantageous to remember or forget previous state, making assessing typically intractable problems with very complex data possible. LSTM has been applied to several temporal processing problems in public health, including influenza,[86] colorectal cancer,[87] and fall prevention.[88] We will build a machine learning model using Keras and TensorFlow in R. First, we will split time-series data into a training set and a validation set of users, in an 80/20 split. We will train a LSTM neural network on the training set. Once trained, we will layer on auxiliary calendar and weather data to try to reduce the model's error and maximize predictive power. We will compare models using several fit indices: Mean Absolute Percentage Error, Root Mean Squared Error, and Root Mean Squared Percentage Error, and plot predicted scores across each model with observed scores to determine closest match. We will apply the model with best fit and most accurate prediction to time-series in validation set to assess how well it predicts new data.

3C.4.3. Integrate Predicted Outdoor Activity with Sun Protection Advice. Algorithms from our *sunZapp* mobile app[71] (see **Box 2**) will be employed to tailor sun safety advice to predicted outdoor physical activity *and personal risk* of *Strava* users. *Machine learning model will predict users' risk of UV-exposure (i.e., timing and duration of outdoor exercise). Users' personal risk (i.e., skin type: hair and eye color and skin tanability[89]) will be collected in SS user profile when first accepted by Strava users. We will use 5-day Planner in sunZapp to combine predicted time and duration of outdoor activity bout, forecast UV Index, and personal risk to calculate Strava users' time to sunburn and provide tailored sun safety advice, based on the World Health Organization's (WHO) Global UV Index guidelines.[90] To be modified based on results from focus groups,* we plan to deliver the advice in the evening and morning before the predicted outdoor physical activity bout (and during extended bouts) through *Strava* comments and emails (see **3C.5**.). *Because Strava users get notified by email and in the app, they likely will see the message before their next outdoor activity.*

Box 2: Prior Research on Sun Safety Advice Algorithms

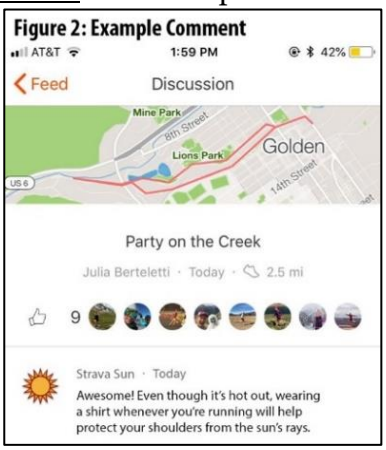
Under a contract from NCI (HHSN261200900025C), Dr. Buller and KB programmers created algorithms to deliver real-time tailored sun safety advice over a mobile app, *sunZapp*.[71] Algorithms first converted hourly UV Index forecast for each 0.5° latitude-longitude grid worldwide published daily by National Oceanic and Atmospheric Administration to UV radiation level. Using skin sun sensitivity input by the user, it calculated time to receive 1 MED of UV (i.e., sunburning dose) to tailor advice on timing of precautions. Algorithms also adjusted UV dose for sunscreen application/reapplication based on sun protection factor (SPF) and for use of shade. Sun protection advice followed WHO's UV Index guidelines.[90] Amount of sunscreen was adjusted for percent of skin uncovered by clothing. In two trials with national adult samples,[59,91] *sunZapp*'s advice improved users' sun safety. In the first trial (n=604 adults from nationwide survey panel), participants assigned to the mobile app improved self-efficacy for sun protection (p=0.022), increased shade use (+7%, p=0.034), and decreased reliance on sunscreen (-6%, p=0.048). Those who used the mobile app minimized time in the sun (+15%, p=010).[59] In the second trial (n=202 adults recruited online), *sunZapp* increased wide-brimmed hat use (+6%, p=0.045). Compared to non-users, young adults who used *sunZapp* wore wide-brimmed hats (+16% 20 year olds, +9% 25 year olds, p=0.03), women employed all sun protection practices (+4%, p=0.04); and more educated individuals spent less time outdoors at midday (-17 hours, p=0.03).[91] An algorithm also delivers predicted UV levels and sun protection advice for up to 5 days from the present time (i.e., 5-day Planner). This routine that will be expanded to provide advice to *Strava* users, tailored to time, location, season, and personal risk, before the next time they are likely to exercise outdoors.

3C.5. Development and Production of Prototype *Strava Sun* (*SS*) Intervention (Specific Aim 3)

The prototype *SS* intervention will be delivered through *Strava*. There are 2 components: 1) a machine learning algorithm that predicts future outdoor activity based on *Strava* users' retrospective data (see **3C.4**) and 2) a

database that delivers tailored, evidence-based sun safety emails/comments through the *Strava* interface based on predicted outside physical activity, forecast UV, and *personal skin cancer risk*. *SS* development will follow KB's standard iterative production. We will develop goals and behavioral objectives and outline of messages based on focus group results, theoretical models (see **3A.3**), *GSS* social media messages (see **Box 3**), and *sunZapp*'s advice algorithms (see **Box 2**). KB media developers will produce a database with the *Strava* API and algorithm predicting future outdoor physical activity (see **3C.4**) to power the *SS* interface and schedule sun safety messages (when and to whom). The *SS* database will determine time of day, date, location (from Wi-Fi, cellular network, or GPS), and forecast UV, and use *sunZapp* algorithm and outdoor physical activity algorithm to provide emails/activity comments over *Strava* platform advising precaution when UV Index is forecast to be ≥3 during predicted outdoor activity time (i.e., risk is moderate, high, or extreme based on Global UV Index standards[90]). Messages will be sent at various times, based on outdoor activity algorithm (see **3C.4.2**); tentatively, we plan to deliver advice on the evening before outdoor activity bout to promote planning for sun safety (e.g., forecast UV Index for time outdoors; packing sunscreen and selecting hats, sunglasses, and protective clothing), just before predicted time of exercising, and during extended exercise (e.g., reapply sunscreen; use shade). *We will ask focus groups about timing preferences (see **3C.3**)*. *Strava* users are accustomed to receiving emails/comments and most check them regularly (see **Box 1**). *SS* messages will be available through any device. Users will not need to interact with a separate app/interface. A specifications document with technical requirements, database plans, flow charts, and content outline will be developed.

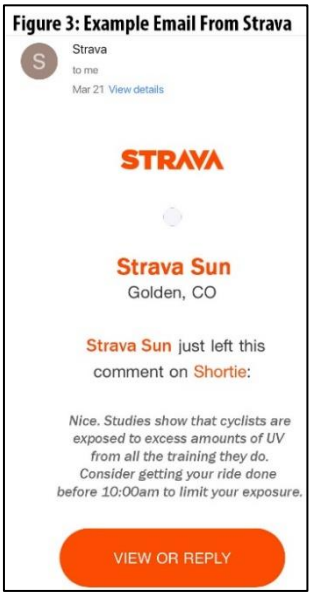
Figure 2: Example Comment



3C.5.1. Utilizing the *Strava* API. With publicly-available tools from *Strava*,[66] KB programmers will create the *SS* database. We will identify the user data (i.e., log dates and time and duration for outdoor physical activity) needed to access and create authentication. Before collecting any data, users will view an authentication screen (see **Figure 1**) presenting the data access terms and conditions (abiding by *Strava*'s developer rules) and information on the study. By accepting the authentication screen, participants give consent for us to access their *Strava* activity data, which KB developers will see through the *Strava* API. Data privacy is extremely important given recent cases in the news and European Union's General Data Protection Regulation.[92] *We will only access Strava data for users who actively give permission.* If a user has accepted, but is ineligible, their *Strava* data will not be used. We will cease data access at the end of the study. Access to the *Strava* API will be informed by Mr. Carmichael, recent *Strava* software engineer and project consultant (see Letter of Support), *who has confirmed the feasibility of the above approach.*

3.C.5.2. *Strava* Comments and Emails. Using *Strava* API, messages will be delivered in emails/ activity comments through the *SS* profile (e.g., "Nice job! Be sure to apply sunscreen and wear a hat for your run tomorrow afternoon. The UV will be high in Denver between 1:00 and 4:00pm."). On all activities, there is a comment box (see **Figure 2**). Depending on the user's settings, *Strava* sends emails to users when a comment is made on their activity (see **Figure 3**), activating a red notification symbol on their *Strava* dashboard, alerting them to the new comment. Users will also receive notifications in the *Strava* mobile app on their phone (e.g., red icon alerts; banner notifications; push notifications), depending on personal settings. Users can reply to the comment and tag the commenter (i.e., *SS*). Additionally, emails will be sent through the *Strava* API ahead of predicted outdoor activity to provide more advice than can be left on a comment (e.g., links to CDC resources or news articles on sun safety). *Strava does not record if an email/comment was read or not, so we will assess users' SS message recall in posttest surveys.*

Figure 3: Example Email From Strava



3C.5.3. Usability Testing of *Strava Sun*. *SS* will be tested for usability with 30 *Strava* users, randomly selected from those who consented to data access for the algorithm (**3C.4.1**) after being invited through the *Strava* API (3 reminders over 4 weeks; compensation=$40). The first 30 to respond will be enrolled and receive the informed consent form, before starting a condensed 4-week *SS* intervention. Next, they will complete a phone interview, assessing message recall, message delivery/ timing *(e.g., if message received before next physical activity bout),* accuracy of outdoor activity predictions, and usability of and problems with *SS* comments/ emails. *They will be asked about sun protection practices during physical activity.* Investigators will review results and programmers will fix problems in *SS*.

Box 3: Preliminary Studies: Sun Safety for Vacationers and Outdoor Recreation

Dr. Buller led a series of studies promoting sun safety in outdoor recreation with the *Go Sun Smart* (*GSS*) intervention. In two randomized trials with the North American ski industry (R01CA81028; R01CA104876; D. Buller, PI), *GSS* increased sun safety practices by ski resort guests,[93,94] i.e., goggles/sunglasses, head

covering, sunscreen reapplication, and overall sun protection (all p<0.001). In a third trial at warm-weather destination resorts in North America (R01CA152411; P. Andersen, PI, D. Buller, Co-I), exposure to *GSS* improved observed sun safety at waterside recreation;[18] self-reported use of any sun protection (i.e., shade, clothing, or sunscreen) improved in the intervention group.[95] *GSS* contained a library of social media messages for Facebook and Twitter feeds which will be used as a basis for the *SS* email and comments.

3C.6. Pilot Field Trial (Specific Aim 4)

The feasibility of the *SS* intervention with adults *Strava* users (n=226) who consent to provide activity data will be tested in a pilot field test using a randomized pretest-posttest controlled design. Participants will be recruited through posts in *Strava* clubs. They will click on a link to learn more about the study and complete a screener (eligibility: 18 or older; reside in the United States; read/write in English; upload outdoor activity to *Strava* ≥3 times per week) and then be provided with a) consent form, b) authentication protocol to permit access to their *Strava* data, and c) pretest survey. Once our target sample is recruited (n=226 *Strava* users), we will disconnect the authentication. Based on our pilot data (**Box 1**), we anticipate that 41% will be high users (upload 5+ days per week) and 43%, medium users (upload 2-4 days per week). Participants will be randomized to either *SS* or untreated control group in a 3.5:1 ratio. Power analysis was estimated using treatment effect sizes from our *sunZapp* trials (**Box 2**). For this small-scale pilot test, we powered on the quasi-experimental pretest-posttest comparison within the *SS* group, using a paired samples t-test. Assuming an effect size of 0.23 in sun protection pre to post and a correlation between the pre/post scores of 0.5, , we will need 181 participants with complete pre-post data to detect an effect of this size with power of 0.80 in a 2-tailed p=0.05 test, which we factored up to 226 to account for 20% attrition. A small control group (n=50) will be included to see if a no-treatment control condition is acceptable to users and estimate follow-up rates for planning a randomized trial. It is not designed to test for statistical differences with the treatment group. *Randomization should assure that the control group is similar on age, gender, and activity level.*

Eight months are allotted for the field trial. Participants will be enrolled on a rolling basis and participate in *SS* for 4 months in the intervention condition. After pretesting, participants will receive sun protection messages from *SS* through emails/comments based on algorithm results from their profile and activity data (see **3C.4**). All participants will complete posttest survey 4 months after randomization. They will receive links to both surveys via email and text and complete them using KB's QuestionPro software. Surveys will contain measures of self-efficacy, response-efficacy, sun protection practices, *sunburn (defined as skin red and/or painful from exposure to the sun in the past four months),*[96] and barriers to sun protection from the *sunZapp* and *GSS* trials (see **Boxes 1** and **2**), recall of *SS* messages (i.e., times they read the notifications; preferred type of message), and acceptability of receiving *SS* messages in future. We will assess physical activity behaviors at both pre- and posttest, using the validated International Physical Activity Questionnaire (IPAQ)[97,98] long form and current activity data uploaded by participants to *Strava*. Intervention participants will be compensated $65 total ($30 pretest, $35 posttest) and control group participants $45 total ($20 pretest, $25 posttest).

3C.7. Data Management and Analysis

Focus groups will be recorded, transcribed, and analyzed using Atlas-ti© software by assigning codes that categorize views, themes, and opinions, rely on constant comparative analysis[99] *and a phenomenological approach.*[100,101] *In the pilot field trial, the primary outcome is change in sun protection pre to post, analyzed with a paired t-test; changes in perceived risk, response- and self-efficacy, and sunburn will be compared pre to post, similarly. Message recall (i.e. receipt of the messages), users' reactions, barriers, and willingness to use SS, will be summarized as proportions/means and compared on demographics (including if treatment differs as a function of biological sex), using chi-square tests, t-tests, and correlations at p=0.05 (2-tailed). We will likewise test if SS messages (message recall) reduced physical activity (IPAQ).* Recruitment rates and duration, and follow-up rates, will be described to plan a future randomized trial on *SS*. Data will be de-identified and stored in KB's secure servers. Dr. Henry will oversee data analysis. SAS will be used for analyses.

Eligibility Criteria

Focus Group Inclusion Criteria:

- 18 years of age or older
- Able to speak and read English
- Consent to participate
- Being a regular user of *Strava* (*Strava* users group only)
- Can attend focus group in Colorado

Focus Group Exclusion Criteria:

- Having a cognitive or visual impairment that would interfere with participating
- Being unwilling to be audio/video recorded
- Having another family member participating in the research

Retrospective Activity Data Partcipants Inclusion Criteria:

- 18 years of age or older
- Able to speak and read English
- Consent to participate
- Being a tenured *Strava* user (member for at least one year with at least 156 activities uploaded)
- United States resident
- Able to provide authentication for data collection

Retrospective Activity Data Partcipants Exclusion Criteria:

- Having a cognitive or visual impairment that would interfere with participating

Usability Testing Inclusion Criteria:

- 18 years of age or older
- Able to speak and read English
- Consent to participate
- Being a tenured *Strava* user (member for at least one year with at least 156 activities uploaded)
- United States resident
- Able to provide authentication for data collection

Usability Testing Exlusion Criteria:

- Having a cognitive or visual impairment that would interfere with participating
- Having another family member participating in the research

Pilot Field Tester Inclusion Criteria

- 18 years of age or older
- Able to speak and read English
- Consent to participate
- Being a regular user of *Strava* (at least 3 uploads per week)
- United States resident

Pilot Field Tester Exclusion Criteria

- Having a cognitive or visual impairment that would interfere with participating
- Having another family member participating in the research
- Having participated in the focus group and usability testing

Age Limits

18-no maximum

Inclusion of Women and Minorities. We anticipate that 50% of the participants will be women. No exclusions for this project will be based on gender.

For the focus groups, to be held in Denver, CO, we will attempt to recruit racial/ethnic minority participants at similar rates to their representation in Denver. According to the U.S. Census, 1.4% of the population is American Indian or Alaskan Native; 3.4% Asian; 0.1% Hawaiian or Pacific Islander; 10.2% African American; 31.8% Hispanic (any race); 4.1% two or more races; and 49% White.[80]

For the data acquisition, usability testing and the pilot field test, we will attempt to recruit racial/ethnic minority participants at similar rates to their representation in the United States. Potential participants will not be excluded based on race or ethnicity. According to the U.S. Census, 1.6% of the population is American

Indian or Alaskan Native; 3% Asian; 0.2% Hawaiian or Pacific Islander; 4.3% African American; 21% Hispanic (any race); 2.5% two or more races; and 67.4% White. The Targeted Enrollment Table represents a composite of these numbers.

During the recruitment process, Dr. Buller and Ms. Berteletti will screen for participants using protocols approved by the Western Institutional Review Board (WIRB), based on gender, race, and ethnicity. If we find that we are over-recruiting a particular gender or racial or ethnic group considering our accrual targets, we will place them on a waiting list until we have recruited an adequate number of targeted participants.

Inclusion Across the Lifespan. Children under 18 will be ineligible for participation. Children under 18 will not be enrolled in the proposed project because this program is aimed at adults and it is not possible to identify parents through the *Strava* interface to obtain parental consent. Adults of any age, who use *Strava* to track their physical activity will be enrolled (non-*Strava* adult users will also be enrolled in the focus group discussions). Age will be self-reported by the participants.

Recruitment and Retention Plan

Recruitment Procedures. Focus group participants will be recruited on a voluntary basis through Facebook and the investigators' networks in Colorado.

We will recruit 1,000 *Strava* users to provide authentication and consent to collect their retrospective (one year of data) activity data through the *Strava* API. These users will be recruited by posting invitations in clubs on *Strava*. The *SS* database will connect to *Strava* through its public Applications Programming Interface (API). This connection allows us to receive data of participants' physical activity from the *Strava* app when they accept our Terms and Conditions and give permission to access their data. Participants will be provided with an online consent form. They will not participate in the intervention or be contacted for follow-up surveys.

Usability testing participants will be randomly selected from those who consented to data access for creation of the outdoor activity algorithm. They will be re-contacted through the *Strava* API by email and comments (with 3 reminders over 4 weeks) and invited to participate in the usability testing. The first 30 to respond to the emails will be enrolled and receive the informed consent form before starting a condensed four-week intervention period. They will be called for a follow-up phone interview after the four-week period.

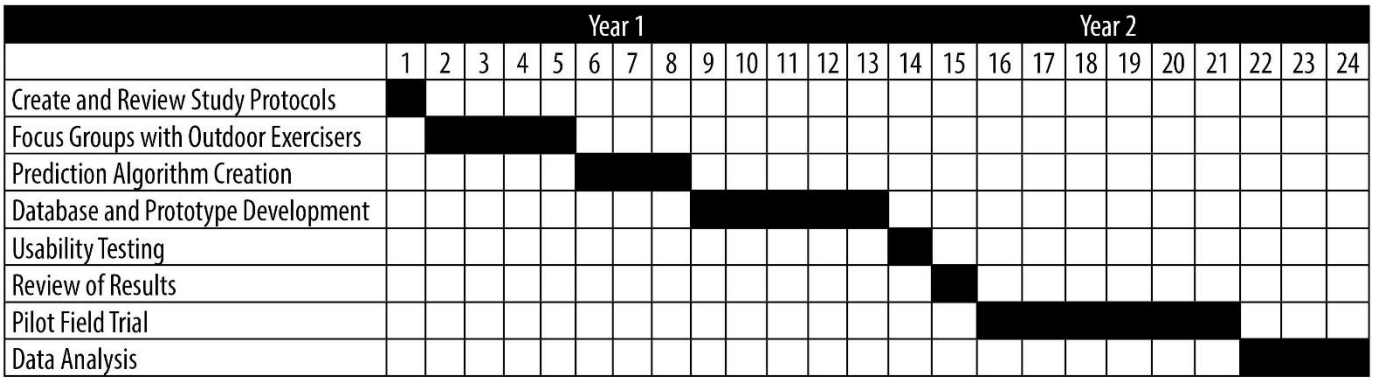
Pilot field test participants will be recruited through postings in *Strava* clubs and inviting participants to click a link to the study website to learn more about the study, where they can complete a brief screener before being provided with an informed consent form, and prompted with the authentication protocol to provide access to their *Strava* data, complete a profile with user characteristics, and complete the pretest survey.

Retention and Adherence. Participant retention and compliance with study procedures are paramount concerns. To address these factors, we have devised and implemented a retention and compliance plan. This plan addresses: 1) how participants are successfully retained throughout the course of the two-year study and 2) how participants are educated about the requirements of the program. Participants may exhibit issues that pose threats to their successful completion of the study, even though they are eligible to participate. These issues are often referred to as "red flags" in the clinical trials literature. Red flags can be circumstances, personal traits, or physical or psychological limitations that could pose potential threats to a candidates' successful retention in a long-term study. Participants may disclose these issues with the study staff publicly or privately, or study staff may suspect these issues after an interaction with a participant. Study staff will be trained on identifying red flags and other barriers to successful participation to help facilitate honest conversations with participants regarding their interest and motivation for being in the study.

To achieve the goal of high participant retention after inclusion into the program, we will use supplemental materials, personal contact, telephone reminders, motivational interviewing techniques, incentives including compensation for assessment completion, and problem solving to overcoming barriers. Participants will be compensated to incentivize and improve study retention. Participants will be compensated as follows: $50 for participation in the focus group (higher to account for travel to KB office for focus group); $40 for participation in the usability testing; and $75 ($35 at baseline and $40 at follow-up) for participation in the pilot field test. Participants who provide access to their retrospective *Strava* activity data for development of the outdoor physical activity algorithm will be entered into a drawing for $50 gift cards (they will not complete any surveys or participate in the intervention). The participants randomized to the no-treatment control condition in the pilot field trial will receive $45 ($20 at pretest; $25 at posttest) for completing the surveys.

2.7. Study Timeline

FIGURE 1: PROJECT TIMELINE



Protection of Human Subjects
RISKS TO HUMAN SUBJECTS

Human Subjects Involvement, Characteristics, and Design. The subject population for this study is adults who engage in routine physical activity and log their activities on *Strava,* an activity tracking and social network site with accompanying mobile app. Both males and females will be included and participants will be of various ethnic and racial backgrounds. Potential participants will not be excluded based on ethnic group or race. Participants will be asked to complete consent and demographic forms. The demographics of the expected participants are detailed in the Targeted Enrollment Table. Human subjects will include:

1. Focus Group Participants (n=32): Adult *Strava* users and nonusers will provide feedback regarding concept of *Strava Sun* (*SS*), preferred content of sun safety advice and delivery methods, their outdoor activities and time of day/day of week they typically exercise outdoors, variation in outdoor exercise by season, confidence and barriers to using *SS* advice, utility/effectiveness of advice for sun protection through *Strava* community, willingness to use the advice while exercising, and concerns around privacy. Current *Strava* users (n=2 focus groups; n=16 total; 8 male, 8 female) will be convened in Denver for a discussion on a) *Strava* use, b) exercise behaviors, c) concerns about privacy, d) willingness to receive recommendations on preventing sunburn while exercising, and *e) use of sun protection while exercising*. A group of outdoor exercisers who do not use *Strava* will also be convened (n=2 focus groups; n=16 total; 8 male, 8 female; *matched to Strava users in focus groups on race, gender, and activity type*). Inclusion criteria are 1) being 18 years of age or older, 2) speaking and reading English, 3) consenting to participate, and 4) being a regular user of *Strava* (*Strava* users group only). Exclusion criteria includes 1) having a cognitive or visual impairment that would interfere with participating, 2) being unwilling to be audio/video recorded, and 3) having another family member participating in the research. Focus group transcripts will be coded and analyzed using ATLAS.ti® software. Participants will be compensated $50 for their time.
2. *Strava* Users in the USA to Provide Activity Data (n=1,000): These participants will not answer survey questions or actively participate in the program. They will provide authentication to collect their retrospective (one year of data) activity data through the *Strava* API. We will not collect their current activity data. The data being collected and what it will be used for will be detailed in the authentication. These users will be recruited by posting in clubs on *Strava*, until 1,000 eligible users have agreed to provide their data. Eligibility requirements include: 1) having at least 156 activities uploaded in one year of *Strava* use; 2) providing authentication. They will be entered to win one of 10 $50 gift cards for allowing access to their data.
3. Usability Test Participants (n=30): Usability testing of the intervention will be conducted with 30 *Strava* users who previously consented to provide their data for analysis and creation of the algorithm (see above). These participants will have already provided consent for data access, but will be re-contacted by email through the *Strava* API to ask if they are interested in participating in the usability testing. The first 30 *Strava* users to respond to the emails will be enrolled and will receive informed consent before starting a condensed intervention period. Identifying the 30 participants in the *SS* database, we will send out the created messages by email and comments for four weeks. At the completion of the four-week period, participants will complete a phone interview with project staff to assess recall of messages, message delivery and timing *(e.g., if message received before next physical activity bout)*, accuracy of outdoor activity predictions, and usability of and problems with *SS* comments/emails. *They will be asked about sun protection practices during physical activity.* Eligibility requirements include: tenured *Strava* user (member for at least one year with 156 activities);

18 or older; United States resident; can read/write in English; and provides authentication for data collection. Participants will be compensated $40 for their time.

4. Pilot Field Testers (n=226): *Strava* users (n=176) will test the prototype *SS* interface while interacting with the *Strava* platform and community in the United States. They will be recruited through posting an invitation in *Strava* clubs. The invitation will link them to the study website where they will receive more information and be screened and consented. Participants will provide feedback on the design, appeal, and functionality of the *SS* message delivery and complete measures of self-efficacy, response-efficacy, sun protection practices, *sunburn (defined as skin red and/or painful from exposure to the sun in the past four months),* and barriers to sun protection, recall of *SS* messages (i.e., how many times a day they read the notifications, how much they liked receiving them, preferred type of message), and acceptability and willingness to use *SS* interface through *Strava* in the future. Additionally, a small control group (n=50) will be included to determine if a no-treatment control condition is acceptable to *Strava* users and estimate follow-up rates for planning a randomized trial. The control group sample is not designed to test for statistical differences in comparison to the treatment group. *Randomization should assure that the control group is similar in age, gender, and activity level.* Inclusion criteria are 1) being 18 years of age or older, 2) writing and reading English, 3) being a regular user of *Strava (*at least 3 uploads per week*)*, 4) living in the United States, and 5) consenting to participate. Exclusion criteria includes 1) having a cognitive or visual impairment that would interfere with participating, 2) having another family member participating in the research, or 3) being unwilling to give consent. They will also be excluded if they participated in a focus group or usability testing. Participants will be compensated $65 ($30 at baseline and $35 at follow-up) for their time. Control group participants will be compensated $45 ($20 pretest, $25 posttest) for completing the surveys.

Study Procedures, Materials, and Potential Risks.

Study Procedures: The two-year R21 research will use scientifically-rigorous formative research and pilot-testing. Investigators will create study protocols in Month 1. Outdoor exercisers, including *Strava* users and non-users, will give input on initial concepts for *Strava Sun* (*SS*) in focus groups during Months 2-5. An algorithm that predicts time of day and day of week spent exercising outdoors will developed using anonymous *Strava* user data in Months 6-8. A prototype of the *SS* database will be developed in Months 9-14 and the messages and delivery methods will be tested for usability and acceptability with *Strava* users. In Month 15, investigators and expert consultant will give input on the formative research results, validity of the time outside algorithm, usability of *SS*, and the pilot-test plans. In Months 16-21, the feasibility of *SS* will be tested in a pilot field trial with *Strava* users. Data analysis on the pilot field trial will be performed in Months 22-24. All procedures will be approved by the Western Institutional Review Board (WIRB), KB's IRB (IRB costs are paid from KB's indirect costs).

The collaborating site is Colorado State University (FWA No. 00000647).

Study Materials: Data from focus groups, usability testing, and pilot field test will be collected in the form of handwritten notes, video and audio tapes, computer files, transcriptions, online surveys, and web server records. We will collect retrospective (one year of data) activity data through the *Strava* API to develop the algorithm predicting outdoor physical activity. Data will be obtained specifically for research purposes and will be kept confidential with access limited to the Multiple Principal Investigators (MPIs), Co-Investigator, specific project staff, and consultants. Data will be obtained specifically for research purposes and will be kept confidential with access limited to the MPIs and specific project staff. All data will be kept in locked files.

Potential Risks: Potential risks to adult *Strava* users and non-users are minimal, mainly social or psychological. Participants may be hesitant or uncomfortable responding to questions about outdoor physical activity and actions to protect their skin from the sun. They also may not feel comfortable responding to questions about their opinions on appeal, ease of understanding, navigation, and functionality of the prototype *SS* and barriers and willingness to use it. Responses will not be shared with anyone except project staff. All participants will be allowed to withdraw at any time if they are uncomfortable with the discussions or survey questions. Responses to the questions will be coded to protect confidentiality, and participants may choose to not answer questions. WIRB requires that the option not to answer be provided to all participants on all questions.

ADEQUACY OF PROTECTION AGAINST RISKS

Informed Consent and Assent. Human Subjects oversight will be conducted by the Western Institutional Review Board (WIRB; DHHS IRB Reg. No. IRB0000053; FWA No. 00003715), KB's IRB. All consenting procedures will be IRB-approved. The WIRB will review and approve subject materials (including

recruitment advertisements, letters, consent forms, survey instruments) and data management procedures prior to study implementation.

Focus group participants will be presented with IRB-approved consent forms that will describe the purpose of the project, risks and benefits, and selection criteria. Participants will be given an opportunity to ask questions and have them answered. For focus groups, consent will occur in-person and will be administered by investigators. For the usability test and for participants providing retrospective data, a link to an online informed consent document will be provided to participants. Participants will have the option to print the consent form. For the pilot field test, online consent will occur prior to the pretest survey and randomization. Again, there will be an option for participants to print the consent language.

Participation will be voluntary, with the right to stop participating at any time. No clinical data will be captured. Responses to survey questions and the discussions would not reasonably place the participants at risk of criminal or civil liability or be damaging to the participants' financial standing, employability, or reputation. All data coding, data entry, analysis and reporting activities will be conducted by KB and CSU. Participants will be informed that they have a right to withdraw from the study at any time. No additional data will be collected from participants who decide to withdraw; however, the information collected prior to withdrawal will remain in the study.

Protection Against Risk. To minimize potential psychological risk, the focus group facilitator (Ms. Berteletti) will establish ground rules for discussion that includes an open, non-evaluative exchange of comments and ideas. All surveys and data collection forms will be approved by WIRB and employed under the supervision of senior project staff. Participants will be told that they do not have to answer any questions which make them uncomfortable and can terminate their participation at any time they wish.

Notes, surveys, audiotapes, and data will be accessible only by research personnel. Responses to survey questions would not reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation. No clinical data will be captured. The data collected on this project will be obtained with the "use of educational tests, survey questions, [and] interview procedures."

Data will be de-identified before being submitted for analysis. All notes and printed data collection forms and questionnaires will be stored in locked cabinets at KB offices; all electronic data files will be stored on KB's secure network servers behind computer firewalls, with routine backup. All identifiers will be stored in secure files and behind computer firewalls and results will be shared in aggregate form only. All human subjects participating in this project will be 18 years and older.

To minimize the risk to participant data privacy and protection through the internet, KB programmers will monitor and maintain all programs and databases and use an established protocol for participant cyber safety and security. KB is HIPAA complaint and maintains vigorous technical safeguards which are built into KB's IT system to protect information and to control access to it. All participant information will be protected using 128-bit SSL encryption. All Internet information must travel through KB's CISCO ASA 5505 firewall using ASA software version 9.0(2) encrypted 3DES-AES. Traffic then moves to KB's web server. This server is a Dell Power Edge T-310 with 32 GB of RAM, 1.5 TB available disk space on a RAID 5 redundant drive system using Windows Hyper-V Operating System. The virtual web server uses Windows 2008 IIS Web server software. User click stream data is collected using Webtrends and is saved to our SQL server. Enrollment data is transferred to our SQL server (Windows Server 2008) which uses MS SQL server 2008 R2 database software. This machine is also a Dell Power Edge T-310 with 32 GB of RAM, 1.5 TB available disk space on a RAID 5 redundant drive system using Windows Hyper-V OS and has no access to the Internet. When a participant is directed to take a survey, their computer is linked to a third server in our system which is also a Dell Power Edge T-310 with 32 GB of RAM, 1.5 TB available disk space on a RAID 5 redundant drive system using Windows Hyper-V OS. This server runs Question Pro Survey software. Data from each survey is collected and saved to the SQL server mentioned above. All network traffic is sent from the CISCO firewall to a Syslog server using Manage Engine Firewall Analyzer 7 software. Login and network activity is required and monitored by Windows servers to an event viewer. Servers are kept in a locked room on site. All files are secured locally using MS NTFS. All traffic between servers and PCs on the LAN are digitally signed communications. Each local machine has a built-in operating system firewall. Data from the Question Pro surveys are exported for data analysis by local machines which are also behind KB's firewall. Access to this exported data by outside sources is done using File Transfer Protocol (FTP) protected by Secure Socket Layer (SSL). SQL databases are protected by the built-in security of MS Active Directory. Data is backed up using a Quantum SuperLoader 3 tape library using LT04 tapes that contain 800 GB/1.6 TB of info (depending on compression used). The device holds 16 tapes. The backup software used by KB is Symantec Back-up Exec. 2012. Daily backups are incremental and tapes are recycled weekly. Weekly backups are done every Friday and are recycled every 6 weeks. Monthly backups are done on the last Friday of each month and are full backups of all data. These monthly tapes are stored in a secure offsite facility and are never recycled.

With regards to app security on personal devices, Apple (e.g., iOS programmed apps) completes a full

security review of any app installed on a phone including beta/testing apps through their App Transport Security (ATS). KB App programming will include all key developer-related features to ensure ATS security. Google Play (Android programming) does not conduct a similar security check; however, KB has extensive experience with developing and marketing both iOS and Android programmed apps using an established protocol (SSL/TLS: Secure Sockets Layer/Transport Security Layer) and has never experienced a data breach. All usability testers will be informed that information being sent to and from their device will remain encrypted and password protected, and at any time they are allowed to opt-out of the program.

POTENTIAL BENEFITS OF THE PROPOSED RESEARCH TO PARTICIPANTS AND OTHERS

Potential benefits for all participants will be the knowledge that they have helped in development of an innovative mobile delivery system for preventing the development of skin cancer. Participants also may increase their knowledge of effective sun protection practices. A sun protection interface for the *Strava* platform will provide individuals who engage in regular physical activity advice to help them plan for and practice sun safety during outdoor activities.

A description of this study will be available on http://www.ClinicalTrials.gov, as required by U.S. Law. This website will not include information that can identify the participant. At most, the website will include a summary of the results. This website can be searched at any time.

IMPORTANCE OF THE KNOWLEDGE TO BE GAINED

Skin cancer is the most preventable cancer because excessive exposure to ultraviolet radiation (UV), the primary risk factor, is easily modifiable. Unfortunately, skin cancer rates continue to rise and many Americans experience sunburns. Physical activity is associated with increased prevalence of sunburn and melanoma. The proposed research is important because it will develop new mHealth strategies to reduce the epidemic of skin cancer in the United States. In addition, the app will provide Americans who engage in physical activity, a healthy lifestyle behavior, real-time advice on sun safety, when and where they need it.

TRAINING IN THE ETHICAL CONDUCT OF RESEARCH WITH HUMAN SUBJECTS

The MPIs, Co-Investigators, and study personnel have completed the required training in the Ethical Conduct of Research with Human Subjects. Necessary documentation will be provided to NIH prior to funds being awarded.

DATA AND SAFETY MONITORING PLAN

A data safety and monitoring plan will be implemented for this R21 project, following well-established procedures at KB. We will not set up an independent Data Safety and Monitoring Board because this study is low risk; does not include use of therapeutics; and physical risks are not anticipated. However, the progress of the formative research will be monitored monthly by the MPIs, Co-Investigators, and study staff who will meet weekly to discuss the activities in the project, participant recruitment, and data collection. Any critical events will be reviewed by the investigators to ensure that project methods have not been intrusive or disruptive. Annual reports on data safety and monitoring are submitted to WIRB. Compliance with data collection protocols will be monitored throughout data collection by the investigators. The investigators are responsible for reporting all adverse events to the IRB. Only Grade 1 events are expected on this trial. Thus, adverse events will be reported annually, as required by DHHS. This report will be reviewed by WIRB. Any action taken by the IRB or project investigators resulting in a temporary or permanent suspension of the trial will be communicated immediately by the contact PI to the NIH program officer.

Statistical Analysis Plan

Focus groups will be recorded, transcribed, and analyzed using Atlas-ti© software by assigning codes that categorize views, themes, and opinions, rely on constant comparative analysis and a phenomenological approach. In the pilot field trial, primary outcome is change in sun protection pre to post, analyzed with a paired t-test; changes in perceived risk, response- and self-efficacy, and sunburn will be compared pre to post, similarly. Message recall (i.e. receipt of the messages), users' reactions, barriers, and willingness to use SS, will be summarized as proportions/means and compared on demographics, using chi-square tests, t-tests, and correlations at p=0.05 (2-tailed). Power analysis was estimated using treatment effect sizes from our sunZapp trials. For this small-scale pilot test, we powered on the quasi-experimental pretest-posttest comparison within the SS group, using a paired samples t-test. Assuming an effect size of 0.23 in sun protection pre to post and a correlation between the pre/post scores of 0.5, we will need 181 participants with complete pre-post data to detect an effect of this size with power of 0.80 in a 2-tailed p=0.05 test, which we factored up to 226 to account for 20% attrition. Recruitment rates and duration, and follow-up rates, will be described for planning a future randomized trial on SS effectiveness. We will test if sun safety messages (SS message recall) reduced physical activity (IPAQ), using correlations. The control group sample is not designed to test for statistical differences in comparison to the treatment group. Data will be de-identified and stored in KB's secure servers. Dr. Henry will oversee data management and analysis. SAS will be used for analyses.