Supplement 1

Study Protocol

Table of Contents

Summary of Changes to the Protocol

After beginning study recruitment but before any participants were randomized or started in an intervention, a decision was made to reduce the number of arms in the trial from three to two by removing the "unconnected arm". This was due to the study timeline. This change was approved by the Institutional Review Boards (IRBs) and the Data and Safety Monitoring Board. The main protocol change was a reduction in the number of arms. No changes were made to the statistical analysis plan other than updating the power calculation for a two-arm study and removing any analysis related to the arm that was removed. Below is the letter to the IRBs describing the changes to the protocol and with updated power calculations:

BE FIT IRB Amendment

November 10, 2015

Dear Members of the IRB,

Thank you for considering our amendment. As requested, we have provided edits to the consent form in attachment 1. Attachment 2 is the updated WTH content and messaging. Attachment 3 is the updated invitation letter. Our updated analysis plan appears below. The revised randomization scheme can be found at the bottom of page 2.

Statistical Analysis

We will start by comparing the characteristics of those who participate in the BE-FIT trial as compared to those who did not, by using data from the most recent examination cycle attended. We will also compare the characteristics of participants using the Moves app as compared to the Fitbit within the participants who participate. Analysis of covariance and logistic regression will be used to assess for differences in the continuous and dichotomous characteristics, respectively, with adjustment for age and sex.

We will conduct an intention-to-treat analysis (all randomized participants) on the primary end-point of proportion of participant-days achieving goal during the 12-week intervention period. Descriptive statistics (sample size, mean, standard deviation, median, quartiles, minimum and maximum) of the primary endpoint will be presented for each group.

Generalized linear mixed-models (via PROC MIXED in SAS) will be used to compare treatments on the primary endpoint using treatment as the fixed effect of interest, team as random effect, and baseline participant-mean daily steps and month fixed effects. An unstructured within-team correlation structure will be assumed. Of interest will be the comparison between the control arm and the intervention.

Data on a given day may be missing if a participant turns off his/her activity tracking device, disables the study team's permission before data is accessed, does not carry the activity tracking device at all, or prematurely withdraws from the study prior to the end of the 12-week treatment period. We will track which days each participant has missing data during the intervention and follow-up period for the two study arms. We will conduct multiple imputation

using five imputations to impute missing data. All analyses will be conducted on each imputed data set and the results will be combined using Rubin's standard rules.[18] The primary analyses will be based on the multiply imputed data. Sensitivity analyses will be performed where the proportion of days achieving goal is calculated based on non-missing days.

Several sensitivity analyses will be conducted to assess the robustness of our findings. First, the model will be further adjusted using fixed effects for activity tracking device (iPhone, Android or Fitbit). Second, the model will also be evaluated using all data and coding missing data (when a step value was not received) as not achieving goal (in contrast to using only collected data). Third, the models will be evaluated using only the collected data (when a step count value was received).

Similar analyses will be conducted for the proportion of participant-days in which the goal was achieved for the 12-week follow-up period and the mean daily steps in the 12-week treatment and follow-up periods. This model will use the multiple imputed data. Several sensitivity analyses will be conducted to assess the robustness of our findings on mean daily steps. First, the model will be further adjusted using fixed effects for activity tracking device. Second, the models will be evaluated using only the collected data (when a step count value was received. Third, because evidence suggests that most step count values less than 1000 are unlikely to represent accurate data capture, we will conduct a sensitivity analysis and exclude those with values less than 1000 excluded from the sample. All analyses will be repeated for the follow-up period of the study. All analyses will be conducted using SAS (SAS Institute, Cary, North Carolina).

Our modified power calculations for enrolling participants in two-person teams are as follows. We estimate that a sample of 86 pairs (172 individuals overall) will ensure 80% power to detect a 0.15 difference in the proportion of participant-days achieving goal between the two groups, accounting for an 8% drop-out rate, a standard deviation in the outcome of 0.30, and an intracluster correlation coefficient of 0.24. Randomization will occur using block sizes of two teams.

Safety outcomes will be evaluated at the end of the 12-week intervention period and the end of the 12-week follow-up period. Our DSMB will review blinded safety data stratified by study arm after the first 50, 100, and 172 participants have completed the first 12-weeks of the trial. We will compare rates of safety outcomes across all 3 trial arms in a blinded fashion using chi-square and t-tests. We will also compare drop-out rates and PA tracking device compliance.

Original Study Protocol

Behavioral Economics Framingham Incentive Trial: A Randomized, Controlled Trial

July 1, 2015

Background:

Physical inactivity is associated with cardiovascular disease (**CVD**) risk factors, including obesity,[1] diabetes,[1] dyslipidemia,[2] and hypertension,[3] and an increased risk for CVD[4] and death.[5] In the Framingham Heart Study (**FHS**) Third Generation (**Gen 3**) and multiethnic Omni 2 cohort participants, objectively measured accelerometry physical activity (**PA**) data showed that moderate to vigorous PA was associated with healthier CVD risk factor profiles.[6] Despite the health benefits of engaging in **PA**, only about half of FHS participants are achieving U.S. recommended aerobic PA guidelines,[7] consistent with the low rates observed in adults nationally.[6] Experts estimate that increasing PA by only 10% would annually save ½ million lives worldwide.[8]

There has been a lot of interest in using "social" or team-based designs to increase PA. Prior evaluations of the FHS found that rates of smoking and obesity were associated with social connectedness within social networks.[9,10] Evidence from other industries has found that interventions using social comparisons feedback to change electricity use were effective, but only if paired with appropriate social approval or disapproval.[11] PA interventions could leverage such social incentives to design interventions that increase peer support, accountability, and unity towards a common goal.[12] In five pilot studies conducted in 2014 at the University of Pennsylvania, we have found that different incentive and feedback designs resulted in differential effectiveness. Our data support the notion that careful testing of alternative social incentive combinations are important to conduct before deciding which intervention to scale more broadly.

The rapidly expanding availability of **mobile technologies** provides a resource effective way to implement social incentive interventions to improve health. Many types of health devices (e.g. glucometers, pedometers) provide individual feedback on performance (e.g. blood sugar or step count). With wireless devices it is technologically feasible to provide relative feedback at periodic intervals. However, utilizing mobile health (**mHealth**) devices is rarely done and could represent a significant opportunity to improve health behavior at low cost.[12] For our study, we will leverage the University of Pennsylvania's NIH-funded *Way to Health* infrastructure. The platform incorporates automated inputs from wireless devices to capture behavior and deliver automated feedback to participants. While conducting the 5 pilot studies noted above, we found that most smartphones and wearable devices were accurate for tracking physical activity data.[13]

We are conducting a **pilot study** and leveraging information acquired through our **Digital Connectedness Survey** to test the deployment of a social incentive intervention using the FHS Offspring, Generation 3, and Omni cohorts to increase PA. We will leverage the strength of FHS by recruiting trios and nuclear families to test whether social connectedness increases PA. We will utilize a **randomized controlled trial design**. We will test **a social incentive intervention strategy using a team-based design in which participants work together to jointly achieve their PA goals, and a social connectedness intervention with teams comprised of connected individuals vs. teams in which individuals are randomly selected.** The **primary outcome** of the intervention pilot study will be the proportion of individual participant-days that the goal is achieved over a 12-week intervention period. Individuals find PA maintenance challenging[14] and PA sustainability has not been studied systematically.[15] As a secondary aim of the intervention study, we will examine PA goal achievement durability for 12 weeks after the intervention ends.

Specific Aims:

Aim 1. To test the **feasibility** of using a team-based social incentive intervention to increase PA in FHS participants**. *H1:*** The demographic and clinical characteristics of individuals who agrees to be randomized are younger and healthier than those who do not agree to participate. ***H2:*** FHS participants who elect the Fitbit in place of the Moves app will differ based on demographic and clinical characteristics.

Aim 2. To evaluate the **effectiveness** of a team-based social incentive intervention to achieve PA goals among **unconnected individuals on a team** versus a **control group of individuals**. ***H3****: A team-based social incentive intervention comprised of unconnected individuals on a team is more effective for achieving PA goals than a control group of individuals.*

Aim 3. To evaluate the **effectiveness** of a team-based social incentive intervention to achieve PA goals among **connected individuals on a team** compared to a **control group of individuals**. ***H4****: A team-based social incentive intervention comprised of connected individuals on a team is more effective for achieving PA goals than a control group of individuals.*

Aim 4. To evaluate the **relative effectiveness** of a team-based social incentive intervention to achieve PA goals among **connected individuals on a team** compared to the same team-based social incentive intervention among **unconnected individuals on a team** for achieving PA goals. ***H5****: A team-based social incentive intervention comprised of connected individuals on a team is more effective for achieving PA goals than a team-based social incentive intervention among unconnected individuals on a team.*

Secondary Aim. To evaluate whether differences in achieving PA goals between study arms during the 12-week intervention are ***sustained*** during the 12-week follow-up period. ***H6:*** *PA levels in the 12-week follow-up will decline for all arms but remain different in a similar fashion to that during the intervention period.*

Our study has the potential to address the major public health problem of sedentary lifestyle, with its consequent increased risks of obesity and CVD risk factor progression. We bring together a multi-disciplinary team with expertise in PA, CVD epidemiology, biostatistics, and behavioral economics. We propose an innovative linking of a classical randomized clinical trial (**RCT**) design with the many strengths of the epidemiological FHS. Our pilot RCT will examine the effectiveness of different approaches to leveraging social incentives to improve health behaviors, an area with considerable promise for increasing the effectiveness of a variety of health interventions. Ultimately our goal is to leverage our pilot study to implement a larger study in the FHS. We also anticipate ultimately linking these results with the extant FHS genetics databases.

Study Design

Study Sample

FHS study members from any of the existing cohorts (Offspring, Gen 3, Omni groups 1 and 2) with an active email address who participated in the FHS Digital Connectedness Survey will be invited to participate. We will target participants with smartphones to identify unrelated and related individuals (related individuals defined as trios [parent, 2 adult children; 3 adult children; spousal pairs plus ≥1 child] who are willing to participate in our pilot. We will also recruit those without a smartphone who have an email address and are willing to use a Fitbit to track their PA.

In order to facilitate randomization of similar individuals to each of the three arms, all participants will be asked to name up to 6 members of their family and 6 friends who are also FHS participants that could join them if they were randomized to a team arm. Participants will also be given the opportunity to indicate 1 FHS family member and 1 FHS friend who they would not want to be paired with if randomized to a team. We defined an individual as "connected" to another if they are related or identified that person as a friend. Our primary goal will be to use connected individuals that are related. To facilitate this goal we will preferentially target large FHS families to participate in our study. However, if we are not able to form enough related groups we will use the secondary definition of friends to form the connected groups. All eligible participants will be paired into teams of 3 members that are connected and then will be randomized to participate individually in the control group, to a team with members that are randomly assigned and not connected to in the unconnected arm, or as the team of three connected members to the connected arm.

Enrollment: Participants will be required to speak English, provide informed consent without substituted judgement, have an email account, and have either a smartphone or daily access to a computer with internet.

Exclusion criteria: Responding yes to at least one of the following questions will result in exclusion from our study:

1) Are you currently participating in any other physical activity studies?
2) Have you been told by a physician that you should not exercise?
3) Are you currently pregnant?
4) Have you had at least one fall with significant injury in the past year?
5) Have you had any surgical procedures in your legs that would prohibit you from being physically active?
6) Do you have any other medical conditions or other reasons why you could not participate in a 6 month physical activity program?

Trial Arms

1) **Control** (Arm 1): individual FHS participants who will not receive the social incentive intervention but will set a daily step goal and receive daily feedback on whether they achieved their goal or not.
2) **Unconnected** (Arm 2): FHS participants placed on a team of 3 randomly assigned and unconnected individuals who will set a daily step goal, receive daily feedback on whether they achieved their goal, and receive the social incentive intervention.
3) **Connected** (Arm 3): FHS participants placed on a team of 3 connected individuals who will set a daily step goal, receive daily feedback on whether they achieved their goal, and receive the social incentive intervention.

Intervention

1) 2-week run-in period and goal setting (Arms 1, 2, and 3)

a. During the 2 week run-in period, we will use the 2nd week of step count data to calculate an individual's mean daily steps ignoring days on which values of less than 1000 steps are recorded (as these likely reflect not using the device rather than lower activity). All

participants are given the option of setting their own daily step goal. We will suggest a goal that is 33%, 40%, or 50% higher than their baseline and inform them that 7000 steps/day is endorsed by the American College of Sports Medicine as being approximately equivalent to federal PA guidelines.[16] However, participants are able to set a goal of their choosing as long as it is at least 1000 steps higher than their baseline step count.

2) Social incentive intervention design (Arms 2 and 3)

a. Participants will be asked to sign a pre-commitment contract to try their best to achieve their daily step goal. Their contract will be available for them to access via the online platform throughout the study.
b. All participants will begin in the bronze level. Higher levels include silver, gold, and platinum.
   - At the end of the study participants in the gold or platinum level will get a modest prize of a coffee mug with the FHS logo.
c. Each week, the team begins with 70 points (10 points for each day)
d. Each day of the week, one member of the team is randomly selected as the representative for the team and the entire team is notified of the selection. If the selected person does not meet his/her step goal on the designated day, the team loses 10 points.
e. Each team member has 5 "lifelines" for which they can request one of his/her team mates to represent the team instead of him/her on the designated day. The lifeline allows individuals that are sick, traveling or cannot reasonably meet their step goal from becoming discouraged and allows them to build upon team support by designating a "lifeline" member to take their place.
f. If the team finishes the week with at least 50 points (i.e. met goal 5 of 7 days which is equivalent to federal PA guidelines[17]) they move up a level (i.e. bronze to silver). If the team does not have at least 50 points, they move down a level. Bronze is the lowest level possible.

3) Outcome

a. **Primary outcome:** proportion of days achieving their step goals over the entire 12-week intervention period.
b. **Secondary outcomes**:
   - Individual participant's mean daily steps during the entire 12-week intervention period adjusted for baseline step count, which is calculated using data from the 2-week run-in period ignoring days with <1000 steps.
   - Proportion of participant-days achieving their step goals over the entire 12-week follow-up period.
   - Mean daily steps over the entire 12-week follow-up period adjusted for the baseline step count.
c. Steps will be tracked using the **Moves smartphone application**, which is free and available for iPhone or Android smartphone platforms. We will also offer a **Fitbit** to participants who do not have either iPhone or Android smartphone platforms but who want to participate in the trial and have an active email account. Both of these devices have been shown to accurately track step counts.[13]

4) Safety endpoints

Participants will have the opportunity to report any adverse events continually through the course of the study. In addition, at the end of the intervention period (12 weeks) and the end of the follow-up period (12 weeks), we will ask participants to fill out a questionnaire that specifically asks whether they have sustained any injuries or hospitalizations. All participants with adverse events will be contacted by the study coordinator and the Adverse Event form will be filled out. The study PI will review these events within 7 days to determine whether any follow-up is needed. Along with safety endpoints, we will track drop-out rates and non-compliance with the PA monitoring devices (proportion of days with step counts <1000).

5) Exit interview

At the end of the intervention period and the end of the follow-up period, we will ask participants to complete a questionnaire on their experiences with the study.

6) Duration

d. 2-week run-in period
e. 12-week intervention
f. 12-week follow-up
g. Total: 26 weeks

7) **Statistical Analysis**

We will start by comparing the characteristics of those who participate in the BE-FIT trial as compared to those who did not, by using data from the most recent examination cycle attended. We will also compare the characteristics of participants using the Moves app as compared to the Fitbit within the participants who participate. Analysis of covariance and logistic regression will be used to assess for differences in the continuous and dichotomous characteristics, respectively, with adjustment for age and sex.

We will conduct an intention-to-treat analysis (all randomized participants) on the primary end-point of proportion of participant-days achieving goal during the 12-week intervention period. Descriptive statistics (sample size, mean, standard deviation, median, quartiles, minimum and maximum) of the primary endpoint will be presented for each group.

Generalized linear mixed-models (via PROC MIXED in SAS) will be used to compare treatments on the primary endpoint using treatment as the fixed effect of interest, team as random effect, and baseline participant-mean daily steps and month the patient was randomized as additional fixed effects. An unstructured within-team correlation structure will be assumed. Of interest are the pairwise comparisons of connected arm vs. individual arm and the unconnected arm vs. the individual arm

Data on a given day may be missing if a participant turns off his/her activity tracking device, disables the study team's permission before data is accessed, does not carry the activity tracking device at all. or prematurely withdraws from the study prior to the end of the 12-week treatment period. We will track which days each participants has missing data during the intervention and follow-up period for all study arms. We will conduct multiple imputation using five imputations to impute missing data. All analyses will be conducted on each imputed data set and the results will be combined using Rubin's standard rules.[18] The primary analyses will be based on the multiply imputed data. Sensitivity analyses will be performed where the proportion of days achieving goal is calculated based on non-missing days.

Several sensitivity analyses will be conducted to assess the robustness of our findings. First, the model will be further adjusted using fixed effects for activity tracking device (iPhone, Android or Fitbit). Second, the model will also be evaluated using all data and coding missing data (when a step value was not received) as not achieving goal (in contrast to using only collected data). Third, the models will be evaluated using only the collected data (when a step count value was received).

Similar analyses will be conducted for the proportion of participant-days in which the goal was achieved for the 12-week follow-up period and the mean daily steps in the 12-week treatment and follow-up periods. This model will use the multiple imputed data. Several sensitivity analyses will be conducted to assess the robustness of our findings on mean daily steps. First, the model will be further adjusted using fixed effects for activity tracking device. Second, the models will be evaluated using only the collected data (when a step count value was received. Third, because evidence suggests that most step count values less than 1000 are unlikely to represent accurate data capture, we will conduct a sensitivity analysis and exclude those with values less than 1000 excluded from the sample. All analyses will be repeated for the follow-up period of the study. All analyses will be conducted using SAS (SAS Institute, Cary, North Carolina).

A priori, we estimated that a sample of at least 414 participants (138 per arm) would ensure 80% power to detect a 15% minimum difference between Arm 1 and Arm 2, and Arm 1 and Arm 3 (2 comparisons), using a conservative Bonferroni adjustment of the Type I error rate using a 2-sided α of 0.025 and assuming an 8% dropout rate. The comparison between Arm 2 and Arm 3 will be a secondary analysis.

Safety outcomes will be evaluated at the end of the 12-week intervention period and the end of the 12-week follow-up period. Our DSMB will review blinded safety data stratified by study arm after the first 100, 200, and 438 participants havev completed the first 12-weeks of the trial. We will compare rates of safety outcomes across all 3 trial arms in a blinded fashion using chi-square and t-tests. We will also compare drop-out rates and PA tracking device compliance.

8) Participant Burden

h. This will be a remote enrollment study. Participants will be contacted via email and will have the option to ask a recruiter questions about the study by email or by phone. Once participants agree to participate, they will obtain an account on the Way to Health platform and complete an online consent and follow the prompts to download the Moves

app. Participants using the Fitbit instead will be mailed their Fitbit and will configure it at home. They will have the option to call the study recruiter for device set-up support. We estimate that enrollment and setup will take no more than **30 minutes**.

i. Participants may choose to receive study communications via email, text message, or interactive voice recording and will receive about 2 daily messages during the 12-week intervention period. An example text message for all groups is “Congratulations! You achieved your daily step goal of X.” In the intervention arms, participants will get a second message related to the intervention: “You have been selected as the representative for your team today. To earn 10 points, please make your step goal of X.” In addition, at the end of each week, team members will receive a message letting them know how the week went. An example message is as follows: “This week, your team achieved its goal 5/7 days. You earn 50 points and advance to the silver level!” We anticipate receiving study communications will take **30 seconds/day per message**.

j. **Total burden:** 30 minute enrollment + 180 messages X 30 seconds each = **120 minutes total over 6 months**

9) DSMB

Although NIH and FDA guidelines do not mandate a DMSB in this setting, we have convened a DSMB for oversight. The DSMB consists of Benjamin M. Scirica MD and Judith Long MD. We will meet with our DSMB members before submission of documents to the IRB, at the end of the intervention period (3 months), and at the end of the follow-up period (6 months).

References

1. Astrup A. Healthy lifestyles in europe: Prevention of obesity and type ii diabetes by diet and physical activity. *Public Health Nutr*. 2001;4:499-515. PMID 11683545
2. Ekelund U, Luan J, Sherar LB, Esliger DW, Griew P, Cooper A, International Children's Accelerometry Database C. Moderate to vigorous physical activity and sedentary time and cardiometabolic risk factors in children and adolescents. *JAMA*. 2012;307:704-712. PMID 22337681 PMCID 3793121
3. Huai P, Xun H, Reilly KH, Wang Y, Ma W, Xi B. Physical activity and risk of hypertension: A meta-analysis of prospective cohort studies. *Hypertension*. 2013;62:1021-1026. PMID 24082054
4. Sattelmair J, Pertman J, Ding EL, Kohl HW, 3rd, Haskell W, Lee IM. Dose response between physical activity and risk of coronary heart disease: A meta-analysis. *Circulation*. 2011;124:789-795. PMID 21810663 PMCID 3158733
5. Nocon M, Hiemann T, Muller-Riemenschneider F, Thalau F, Roll S, Willich SN. Association of physical activity with all-cause and cardiovascular mortality: A systematic review and meta-analysis. *EurJCardiovasc PrevRehabil*. 2008;15:239-246. PMID 18525377
6. Glazer NL, Lyass A, Esliger DW, Blease SJ, Freedson PS, Massaro JM, Murabito JM, Vasan RS. Sustained and shorter bouts of physical activity are related to cardiovascular health. *Med Sci Sports Exerc*. 2013;45:109-115. PMID 22895372
7. Centers for Disease C, Prevention. Adult participation in aerobic and muscle-strengthening physical activities--united states, 2011. *MMWR Morb Mortal Wkly Rep*. 2013;62:326-330. PMID 23636025

8. Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, Lancet Physical Activity Series Working G. Effect of physical inactivity on major non-communicable diseases worldwide: An analysis of burden of disease and life expectancy. *Lancet*. 2012;380:219-229. PMID 22818936 PMCID 3645500
9. Christakis NA, Fowler JH. The collective dynamics of smoking in a large social network. *N Engl J Med*. 2008;358:2249-2258. PMID 18499567 PMCID 2822344
10. Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. *N Engl J Med*. 2007;357:370-379. PMID 17652652
11. Schultz PW, Nolan JM, Cialdini RB, Goldstein NJ, Griskevicius V. The constructive, destructive, and reconstructive power of social norms. *Psychol Sci*. 2007;18:429-434. PMID 17576283
12. Patel MS, Asch DA, Volpp KG. Wearable devices as facilitators, not drivers, of health behavior change. *JAMA*. 2015. PMID 25569175
13. Case MA, Burwick HA, Volpp KG, Patel MS. Accuracy of smartphone applications and wearable devices for tracking physical activity data. *JAMA*. 2015;313:625-626. PMID 25668268
14. Mitchell MS, Goodman JM, Alter DA, John LK, Oh PI, Pakosh MT, Faulkner GE. Financial incentives for exercise adherence in adults: Systematic review and meta-analysis. *Am J Prev Med*. 2013;45:658-667. PMID 24139781
15. Blackman KC, Zoellner J, Berrey LM, Alexander R, Fanning J, Hill JL, Estabrooks PA. Assessing the internal and external validity of mobile health physical activity promotion interventions: A systematic literature review using the re-aim framework. *J Med Internet Res*. 2013;15:e224. PMID 24095951 PMCID 3806547
16. Garber CE, Blissmer B, Deschenes MR, Franklin BA, Lamonte MJ, Lee IM, Nieman DC, Swain DP, American College of Sports M. American college of sports medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: Guidance for prescribing exercise. *Med Sci Sports Exerc*. 2011;43:1334-1359. PMID 21694556
17. 2008 physical activity guidelines for americans. 2008. http://www.health.gov/paguidelines/ (accessed March 10, 2015.
18. Rubin, D. Multiple imputation for nonresponse in surveys. New York: J. Wiley & Sons, 1987.