Feasibility, preliminary efficacy, and accessibility of a twitter-based social support group vs Fitbit only to decrease sedentary behavior in women

M.A. Oppezzo a,*, J.A. Tremmel b, K. Kapphahn c, M. Desai c, M. Baiocchi d, M. Sanders e, J. J. Prochaska a

a Department of Medicine, Stanford Prevention Research Center, Stanford University School of Medicine, Stanford, CA, United States of America
b Interventional Cardiology, Women’s Heart Health at Stanford, Stanford, CA, United States of America
c Epidemiology and Population Health, Stanford University School of Medicine, Stanford, CA, United States of America
d Interventional Cardiology, Women’s Heart Health at Stanford, Stanford, CA, United States of America
e Quantitative Science Unit, Stanford University School of Medicine, Stanford, CA, United States of America

ABSTRACT

Background: Health behavior change interventions delivered by social media allow for real-time, dynamic interaction, peer social support, and experimenter-provided content.

Aims: We tested the feasibility, acceptability, and preliminary efficacy of a novel Twitter-based walking break intervention with daily behavior change strategies and prompts for social support, combined with a Fitbit, vs. Fitbit alone.

Methods: In a 2-group pilot, 45 sedentary women from a heart clinic were randomized to Twitter + Fitbit activity tracker (Tweet4Wellness, n = 23) or Fitbit-only (control, n = 22). All received a Fitbit and 13 weeks of tailored weekly step goals. Tweet4Wellness consisted of a private Twitter support group, with daily automated behavior change “tweets” informed by behavior change theory, and encouragement to communicate within the group. Feasibility outcomes included recruitment and enrollment numbers, implementation challenges, and number and type of help requests from participants throughout the study period. Preliminary efficacy outcomes provided by Fitbit data were sedentary minutes, number of hours with >250 steps, maximum sitting bout, weighted sedentary median bout length, total steps, intensity minutes (>3.0 METS), and ratio of time spent sitting-to-moving. Acceptability outcomes included level of Twitter participation within Tweet4Wellness, and Likert scale plus open-ended survey questions on enjoyment and perceived effectiveness of intervention components. Survey data on acceptability of the features of the intervention were collected at 13 weeks (end-of-treatment [EOT]) and 22 weeks (follow-up).

Results: The study was feasible, with addressable implementation challenges. Tweet4Wellness participants changed significantly from baseline to EOT relative to control participants on number of active hours p = .018, total steps p = .028, and ratio of sitting-to-moving, p = .014. Only sitting-to-moving was significant at follow-up (p = .047). Among Tweet4Wellness participants, each tweet sent during treatment was associated with a 0.11 increase in active hours per day (p = .04) and a 292-step increase per day (p < .001). Tweet4Wellness participants averaged 54.8 (SD = 35.4) tweets, totaling 1304 tweets, and reported liking the accountability and peer support provided by the intervention.

Conclusion: A Twitter-delivered intervention for promoting physical activity among inactive women from a heart clinic was feasible, acceptable, and demonstrated preliminary efficacy in increasing daily active hours, daily total steps, and the ratio of sitting-to-moving from pre to post for the intervention compared with the control. Lessons learned from this pilot suggest that the next study should expand the recruitment pool, refine the intervention to increase group engagement, and select active hours, total steps, and ratio of sitting-to-movement as primary sedentary behavior measures.

1. Introduction

Sedentary behavior is a major, independent risk factor for chronic diseases like heart disease, and early mortality, distinct from lack of aerobic exercise (Matthews et al., 2008; Ward et al., n.d.; Machado de Rezende et al., 2014; CDC, 2020; Lavie et al., 2019; Evenson et al., 2014; Same et al., 2015). The recent physical activity guidelines have added recommendations to reduce sedentary behavior (Same et al., 2015), rendering it an important, though currently understudied, priority for practitioners and researchers.

While epidemiological evidence supports the associations of cumulative sedentary time with disease, the pattern in which the sedentary time is accumulated is also important (Diaz et al., 2017). Prolonged sitting has both acute and chronic deleterious effects on health. Interrupting prolonged sitting breaks, with even a few minutes of walking, can reverse some of these negative effects. While the ideal pattern of sedentary time, including how often and how long to break up sitting, is unspecified in the literature or national recommendations (Giurgiu

* Corresponding author at: Department of Medicine, Stanford Prevention Research Center, Stanford University, Stanford, CA, United States of America.
E-mail address: moppezzo@stanford.edu (M.A. Oppezzo).

https://doi.org/10.1016/j.invent.2021.100426
Received 18 March 2021; Received in revised form 23 June 2021; Accepted 4 July 2021
Available online 6 July 2021
2214-7829/Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
et al., 2020), investigating ways to decrease sedentary behavior in general is important. Because sedentary behavior takes place across many contexts, for example during work, leisure activities, or transportation, and is ongoing throughout the day, behavior change interventions may require a different approach than those targeting an increase in exercise (a single bout activity).

Mobile health (mHealth) behavioral interventions are accessible at multiple timepoints throughout the day and may show a particular benefit for targeting sedentary behaviors (Maher and Conroy, 2016). Reviews of mHealth show mixed effects for sedentary behavior reduction (Stephenson et al., 2017; Fiedler et al., 2020; Buckingham et al., 2019). The inconsistent results are likely in part due to incomparability across studies with different designs (e.g., single vs multiple component interventions), populations (e.g., older vs younger adults), durations (e.g., 3 months to 6 months), intervention targets (e.g., sedentary behavior alone vs sedentary behavior and physical activity together), measure types (e.g., device vs survey measures), and sedentary behavior outcomes (e.g., total sedentary minutes vs prolonged sitting bouts) (Stephenson et al., 2017; Fiedler et al., 2020). Additionally, sedentary behavior interventions often define a prolonged sitting bout differently (e.g., bouts lasting 30 min or more in Shrestha et al. (2018), and 45 min or more in Boerema et al. (2019). However, while variability is high, there is still reasonable evidence for mHealth as a feasible, acceptable, and effective tool to explore for long term sedentary behavior reduction using a modality that allows throughout-the-day intervention access (Stephenson et al., 2017; Fiedler et al., 2020; Buckingham et al., 2019).

Social media, or web apps that allow users to receive, generate, react to, and share content via a social network, are a specific type of mHealth platform harnessed in health interventions, with modest effectiveness (Maher et al., 2016). Not only can social media deliver in-context health information with broad reach, customizability, and easy access (King et al., 2013; Riley et al., 2011); it can also allow users to react, add to the content, and provide social support to other users (an evidence-based behavior change technique (Maher et al., 2016; Pechmann et al., 2015; Bandura, 2001; Davis et al., 2015)). Twitter is a choice social media intervention platform, with high prevalence of use; 73% of the adults in the United States use social media sites, and the majority use these sites daily (NW 1615 L, 2013), often accessed via mobile devices (Twitter by the Numbers, 2020). Twitter has the capability of allowing for private groups to be created that are protected from the public, making it ideal for delivering and privatizing a research intervention. Additionally, Twitter messages (called tweets) have a 280-character limit, which enables messages to be short and accessible. Often used as a supplementary aid to in-person interventions or interventions hosted in other platforms, the potential for utilizing Twitter as a stand-alone platform to deliver health behavior interventions is not yet fully realized (Maher et al., 2016). When used, engagement strongly predicts the benefits (An et al., 2008; Cole-Lewis et al., 2019). Tweet2Quit, a Twitter-based intervention for smoking cessation, was among the first successful interventions utilizing Twitter as the main feature, showing sustained long-term engagement and maintenance of changed behavior compared with controls (Pechmann et al., 2015; Pechmann et al., 2017).

Tweet4Wellness is a Twitter-based platform, modeled after the successful smoking cessation intervention, Tweet2Quit (Pechmann et al., 2015; Pechmann et al., 2017). Within a private, study-created social media support group, Tweet4Wellness delivers daily messaging of evidence-based behavior change strategies, with prompts for group interaction. For the current pilot, the messaging was directed at increasing physical activity and frequent walking breaks.

The primary aims of the current study were to test for feasibility, acceptability and preliminary efficacy of Tweet4Wellness + Fitbit intervention compared with an active Fitbit-only control group in women patients of a heart clinic who did not meet physical activity guidelines. The goal was to identify points for refinement of the intervention, deployment, and the outcome measures to then optimize the treatment for a fully powered test of effectiveness. This qualifies our pilot as stage 1b: intervention development/refinement of the NIH's Stage Model for Behavioral Intervention Development (Onken et al., 2014). Intervention participants received 13 weeks of daily, theory-based behavior change strategies delivered via a private, study-created social-support group in Twitter. As a pilot, we were not designed to detect significant effects. However, we hypothesized that the women who were in the Tweet4Wellness + Fitbit intervention would exhibit greater reductions in sedentary behavior over the 13-week intervention and at the 22-week follow-up compared with Fitbit-only control.

2. Methods

2.1. Study overview

The current study tests the feasibility, acceptability, and preliminary efficacy of a Twitter-based behavior change intervention to decrease sedentary behavior. Full methods and protocol design have been described in detail in the protocol study (Oppezzo et al., 2020).

2.2. Study recruitment, eligibility criteria, randomization

Recruitment emails were securely sent from the University Research Repository System to women in the Women's Heart Health Clinic or Cardiovascular Clinics at Stanford. Eligibility criteria included having a smart phone with an unlimited texting plan, being familiar with social media, not meeting the National Physical Activity Guidelines (assessed by the International Physical Activity Questionnaire (Lee et al., 2011)), and being physically able to safely walk (assessed by the Physical Activity Readiness Questionnaire (Thomas et al., 1992)).

A mandatory in-person or Zoom-remote information session shown to raise retention and increase participant partnership was conducted to review the consent form, study procedures, and research method principles (e.g. why it is important to still participate even if the intervention does not work) (Goldberg and Kiernan, 2005). Of 68 attendees, 45 consented, and were randomized in sets of 10 (final set of 15) to either the treatment (n = 23) or control (n = 22) group, stratified by their baseline self-reported physical activity minutes per week. Specifically, after the first 10 participants consented, they were rank ordered according to their baseline self-reported physical activity. Next, the most physically active participant was randomly assigned to a condition based on a random.org coin flip. The second most physically active person would then be automatically assigned to the opposite group. This continued for the remaining participants in the set of 10.

2.3. Study timeline

Baseline survey data were collected before randomization. After randomization, participant Twitter and/or Fitbit accounts and apps were set up via phone visit with study staff. Baseline Fitbit data were collected for 1 full week prior to the official study start. The Tweet4Wellness intervention and Fitbit weekly step goals were stopped at 13 weeks, and surveys and movement data were collected at end-of-treatment (EOT); additionally, follow up data were collected at 22 weeks (see Fig. 2 for timeline).

2.4. Intervention

Both the treatment (Tweet4Wellness + Fitbit) and control (Fitbit-only) group received the Fitbit self-monitoring component (described below). The treatment group also received the Twitter-based Tweet4Wellness intervention.

2.4.1. Fitbit self-monitoring component

All participants received a study-provided wrist-worn Fitbit Inspire and study-provided Fitbit account connected to Fitabase, a web-based analytics and data aggregation system (Small Steps Labs). The Fitbit
allowed for self-monitoring of daily steps and number of active hours, or number of hours with more than 250 steps. The Fitbit automatically
vibrates at 10 min to every hour between 9 AM to 7 PM if the wearer
does not have 250 steps for that hour, or was sedentary. Participants
were encouraged to open the Fitbit app daily to monitor their activities
and to sync their data with Fitbase. Syncing the data not only aided
data collection, but days synced served as an adherence measure for all
participants. Participants would receive email reminders to sync if they
hadn’t synced for over 24 h. Additionally, all study participants were
texted automatically-generated, personalized weekly text goals for the
week (10% more steps per day, given their average step count the pre-
vious week. Weekly texts stopped after 13 weeks of active treatment, but
Fitbit data were still collected during the follow-up period.

2.4.2. Tweet4Wellness component

To preserve anonymity, treatment participants received a study-
provided Twitter account, and were placed in a private Twitter group.
Automated daily prompts suggesting a behavior change strategy and
couraging group sharing were sent at 9 AM to the group. The strate-
gies were informed by theories of behavior change, namely, Bandura’s
Social Cognitive Theory (Bandura, 2001), Prochaska’s Transtheoretical
Model of Behavior Change (Prochaska and Velicer, 1997; Prochaska and
DiClemente, 1986), Dweck’s Implicit Theories model (Dweck, 1996),
and Gollwitzer’s implementation intentions (Gollwitzer, 1999). These
280-character or less strategy messages were developed by a behavioral
scientist with the help from study team members, pilot tested with
women in the heart clinic and on MTurk, and went through multiple
rounds of iteration and refinement. More detail on these messages can
be found in the protocol paper (Oppezzo et al., 2020). Each strategy was
based in one or more of the above listed theories of behavior change, and
they were grouped according to whether they suggested a change to
thinking (internal) or a change to the surrounding environment
(external) to help support physical activity. Each week of the treatment
alternated between internal or external strategies (e.g. 1 week of 7 daily
internal strategies was then followed by 1 week of 7 daily external
strategies). Treatment group participants also received reminder texts
providing encouragement to engage with the group, based on tweeting
behavior on the prior day. If the participant had tweeted, they were
thanked as reinforcement, if they hadn’t, they were gently encouraged.
The schedule of engagement encouragement is reported elsewhere
(Oppezzo et al., 2020), but ranged from daily to every five days for non-
tweeting throughout the 13 weeks.

2.5. Measures

2.5.1. Feasibility

To determine the feasibility, we tracked recruitment outcomes,
implementation challenges, and the number and type of help requests
throughout the study.

2.5.2. Preliminary efficacy

To measure characteristics of sedentary behavior and physical ac-
tivity, we chose outcomes in line with Byrom et al.’s (2016) guidelines in
addition to 3 outcomes that were visible to all participants on their Fitbit
app as part of self-monitoring. While sedentary behavior can be under-
stood in terms of total sedentary time spent, the pattern of sedentary
behavior is also important. To understand both the amount and the
pattern, the following characteristics, assessed via Fitbit, were exam-
ined: 1. Total sedentary time (minutes/day), 2. Number of active hours
(hours with 250 steps or more in them), 3. Maximum sedentary bout
length (daily non-sleep, longest sitting bout in minutes), 4. Weighted median
sedentary bout length (daily non-sleep prolonged sitting bout length that captures 50% of all sedentary time), 5. Total steps
(steps/day proxying total movement time), 6. Intensity minutes
(minutes/day where activity was 3.0 or more METS), and 7. Sitting-to-
moving ratio (minutes of non-sleep non-movement to minutes of
movement). These are each described in Table 1. Because there is no
single characteristic of sedentary behavior and relationship to move-
ment throughout the day, and given this is a pilot study, we chose these
7 characteristics for behavioral outcomes. While some outcomes are
related to each other, individually they each capture a unique compo-
nent of sedentary behavior and physical activity. Intensity minutes were
the only outcome not explicitly encouraged by the intervention; we
included it to explore whether exercise changed for either group over
the study period.

Table 1

| Variable (unit) | Description | Calculation |
|----------------|-------------|-------------|
| Total sedentary time (minutes) | Total number of waking minutes in a day spent sedentary. | Non-sleep minutes with zero steps |
| Number of active hours (count) | Number of waking hours in the day with at least 250 steps, ~2 min of walking, tracked and visualized on the Fitbit app. | Non-sleep hours with at least 250 steps |
| Maximum sedentary bout length (minutes) | Longest continuous sedentary bout while not sleeping. | Longest period of contiguous minutes of zero steps that is not sleep |
| Weighted median sedentary bout length (minutes) (Byrom et al., 2016) | Length of the sedentary bout that corresponds to the ordered bout length at the median of the total length of bouts. To illustrate, if a day included 2 5-min bouts, one 10-min bout, and 2 15-min bouts, then the total time is 55 min, the midpoint of total time is 22.5 min, and when ordered from shortest to longest (5, 5, 10, 15, 20), the 15 min bout will contain the 22.5th minute, so the weighted median sedentary bout length is 15. | At day level, sedentary bout lengths ordered from shortest to longest to calculate the midpoint where half the sum of all sedentary bouts is captured. The sedentary bout length which contains the midpoint is the weighted median sedentary bout length. |
| Total daily steps (count) | Any movement activity accrued throughout the day. | Total steps taken on days where Fitbit was worn (non-wear days = total steps < 300) |
| Intensity minutes (minutes) | Sum number of minutes spent at either “active” or “very active” (defined by Fitbit) intensities throughout the day. Active and very active are determined via proprietary algorithm using both steps and cadence (or speed of the steps). “Active” minutes are counted if an estimated 3 METS, or Metabolic Equivalents, or higher are being used (or at least 3 times the energy spent than at rest). “Very active” minutes are counted for estimates of 10 METS or more (per Fitbit.com). | Total active and very active minutes |
| Sitting-to-Moving (ratio) | Ratio of time spent sedentary to time spent moving. | Sedentary minutes divided by total minutes where one or more step occurs |

- Indicates self-monitoring variable that is visible to the participant via the Fitbit device and app.
Sedentary behavior and physical activity characteristics were measured daily by the Fitbit device. Fitbit data collection was initiated one week before the Twitter intervention began. Originally, accelerometers collected data at baseline, 13 weeks, and 22 weeks for 7 days, however the device batteries failed and Fitbit data was utilized instead.

2.5.3. Acceptability
To assess acceptability, we measured level of participation within the Twitter support group (number of tweets sent), adherence to syncing the Fitbit (ratio of the number of days Fitbit was synced to all possible days), Likert scale survey questions on enjoyment and perceived effectiveness of the Fitbit and Twitter intervention components, and open-ended feedback on Twitter intervention.

2.6. Data cleaning
Minute level step and sleep Fitbit data were exported from Fitabase. Sleep minutes were excluded from sedentary minute calculations, and days with ≤300 steps were deemed non-wear days and excluded from analyses.

2.7. Data analyses
To model each of the sedentary behavior characteristics as a function of time point, we used mixed-effects models with a random intercept of individual. The main independent variables of interest are condition (treatment vs control) time (baseline vs post-intervention vs follow-up), and their interaction. The models adjust for device usage (syncing of the Fitbit) as a covariate. Specifically, this model was used to test whether the baseline-EOT-follow-up change was different between the intervention and control groups (i.e. time × condition interaction). Each time point for these analyses included several days of Fitbit data from three different time periods: before the Twitter intervention, immediately after, and then at follow-up. Secondary analyses tested the impact of the Twitter treatment engagement on the physical activity outcomes. To explore whether engagement via tweeting mattered within the Twitter group, we ran the same models only for those in the treatment group.

2.8. Data summaries
Participant-stated issues are depicted in Table 4.

3. Results

3.1. Baseline sample characteristics
Demographics and other baseline characteristics for the treatment (n = 23) and control (n = 22) group are described in Table 2. At baseline, compared to the control group, the treatment group reported significantly more sedentary hours per day in the sedentary behavior questionnaire (SBQ; Rosenberg et al., 2010), F(1,43) = 4.1, p = .049, and had a significantly higher BMI, F(1,42) = 9.0, p = .005. All other comparisons by group at baseline were nonsignificant. Of note, while the average age of the sample was 60.6 (12.3), the age ranged between 36 and 87 years old.

3.2. Feasibility outcomes

3.2.1. Recruitment
Recruitment was via the University Research Repository, an “opt-in” list for patients of the hospital to receive information about university research studies. Using this tool, we identified women who were referred to or current patients of the Women's Heart Clinic or Cardiovascular Clinics that had agreed to be contacted regarding research studies. Recruitment details are shown in the CONSORT diagram in Fig. 2. Of 1531 recruitment emails sent, 858 women clicked on the screening survey link, 464 completed the screening, 68 women met the inclusion criteria and attended the information session, and 45 ultimately consented and enrolled. The original target of recruiting 17–25 women per group was therefore met with this first round of recruitment. The relatively small yield for the original email suggests additional recruitment methods and a broader population pool for the next pilot, and opening up to all genders rather than limiting to females.

3.2.2. Implementation challenges
An important outcome of conducting feasibility trials is identifying and solving unforeseen implementation challenges; in e-health studies, this often is related to technology. Feasibility of our intervention does not hinge on the absence of problems, but rather the ability to address them for the larger trial. We encountered the following implementation challenges: accelerometer device failure; Twitter-account shut down; unsupportive social media communication; and frequency of contact. These are outlined in terms of challenge and resolution in Table 3. Participant-stated issues are depicted in Table 4.

3.3. Preliminary efficacy outcomes
Results from baseline to EOT, the treatment group outperformed the control group on several outcomes. For the number of active hours per day, the treatment participants gained an average of 0.19 number of hours [95% CI = 0.39, 0.77], while the control participants lost on average 0.77 h [95% CI = −1.3, −0.23], p = .018. For the average number of total daily steps, the treatment participants increased a mean of 820 steps [95% CI = −3.83, 1724.9], while the control participants decreased on average 570.9 steps [95% CI = −1414.7, 272.9] p = .028. Finally, for the sitting-to-moving ratio, the treatment participants decreased by a mean of 2.0 [95% CI = −4.37, 0.36], while the control participants increased on average 2.1 [95% CI = 0.15, 4.31], p = .014. (See Table 5.) Only the difference in change for the ratio of sitting:moving remained significant at the 22-week follow up. While there are benefits to using a ratio, changes in ratios are difficult to interpret, as they can either result from changes in the numerator (sitting minutes) or the denominator (moving minutes). For this outcome, a decrease in ratio broadly means that the
5

relative relation of sitting to moving improved. A more interpretable measure of sedentary behavior that controls for amount of time spent sleeping is the percentage of awake minutes spent sitting. This measure will be utilized in the next study. There was not sufficient evidence to indicate differences in change in other outcomes between the groups (i.e., total sedentary minutes, maximum sedentary bout, weighted median sedentary bout, and intensity minutes).

3.3.1. Secondary analyses
Within the Tweet4Wellness group over the treatment period, engagement as measured by number of tweets sent had a significant association with the number of active hours (each additional tweet, a 0.11 increase [95% CI 0.01, 0.22] in number of active hours per day, p = .04); total daily steps (each additional tweet, a 289.9 step increase [95%CI 148.3, 431.5], p < .001); and intensity minutes (each additional tweet, a 2.9 intensity minute increase per day [95%CI 1.7, 4.1], p < .001).

3.4. Acceptability outcomes
Through 13 weeks of the treatment phase, only 1 participant discontinued due to a personal reason unrelated to the study. EOT retention was high for both treatment and control groups: 40/45 (88.9%) of participants completed EOT surveys (see Fig. 1). At 22 weeks follow-up, 39/45 (86.7%) completed follow-up surveys. Acceptability of study procedures was good for the study sample overall (~75% slightly to strongly agree). Fig. 3 shows results from the EOT survey questions about self-reported change since baseline. In response to self-perceived progress on self-chosen sedentary behavior or activity goals, 12/21 (57%) control participants and 15/19 (79%) treatment participants reported some or total progress. Engagement, as measured by percent of the condition group syncing Fitbit data on any given day, are depicted in Fig. 4, with higher engagement at the start of the study, but averaging 95% over the study period. The mean (SD) days of any wear with Fitbit (or days with a minimum of 300 steps) was 141.8 (18.25) days for the control group and 129.1 (38.3) for the Twitter group out of 154 days of the trial and follow-up phase.

Treatment-specific feedback was collected from the Tweet4Wellness treatment group only. Tweet engagement with the group decreased over time, but every participant tweeted at least once, with a total of 1304 tweets over the 13 weeks. Fig. 5 shows the proportion of the group who sent at least one tweet throughout the treatment period. Participants sent on average 54.8 (SD = 35.4) tweets on 35.6 (SD = 21.6) different days over the 13-week intervention. Proportions of those who evaluated components of the study as between somewhat helpful to extremely helpful were 65% for the daily prompts, 50% for reading the group posts, 45% for pesting to the group themselves, and 42% for group interactions in general. Tables 6a–6b show sample quotes from the users answering, “what did you like?” and “what would you change?” questions, respectively.

4. Discussion
This pilot feasibility, acceptability, and preliminary efficacy trial tested the effects of Tweet4Wellness, a privatized, Twitter-based support group with daily study-delivered, theoretically-based behavior change strategies to decrease sedentary behavior. Several implementation challenges led to a tightening of the protocol that will highly benefit the next pilot. Compared with a control group receiving the Fitbit only component, the treatment group showed more improvement from baseline after 13 weeks in number of active hours, daily total steps, and the proportion of waking minutes spent sitting to waking minutes spent moving; however, these changes were not maintained at follow-up. There was no significant change on other sedentary behavior characteristics (total sedentary time, maximum sedentary bout, weighted median sedentary bout, and intensity minutes). Within the treatment group, the intervention showed positive uptake, with engagement (i.e., tweeting) having a small, but significant effect on the impact of the intervention on daily total active hours, total steps, and intensity minutes.

This is one of the first experimental studies targeting sedentary behavior to examine Twitter as both a delivery mechanism for behavior change strategies and a venue for social support. Tweet2Quit, the parent study protocol that targeted smoking cessation, showed an association between engagement via tweeting and abstinence, with each additional 10 tweets increasing abstinence by 20% on average (Pechmann et al., 2017). Here, we find a small but significant effect of tweeting on the primary indicator of movement throughout the day, with each additional 10 tweets increasing the # of active hours over the 13-week treatment period by 1.1 h. While only correlational, this suggests that increasing engagement may influence the primary outcome. Given the

Table 3
Implementation challenges and resolutions.

| Implementation challenge | Resolution |
|--------------------------|------------|
| Widespread battery failure of accelerometers leading to 90% of baseline data loss; complaints of accelerometer bulkiness, and neglecting to return the accelerometer led to noncompliance | Fitbit data were analyzed and used for the preliminary efficacy data. |
| Participants began wearing Fitbits on the planned date of study start; at 4:30 AM the morning of planned study start, an estimated 19 of the 23 study-set-up Twitter accounts were permanently shut down due to increased site restrictions around multiple account set-ups. For privacy and original protocol consent reasons, and to maintain study-created Twitter-accounts, we had to postpone the study start by one week to solve the issue | Purchase of 23 email addresses to create new Twitter accounts; set-up of new Twitter accounts with participants in the treatment group; both groups wore Fitbits without direct study intervention for one week providing new baseline data; started the study intervention a week later than originally planned. |
| Unsatisfactory social media communication; participant repeatedly posted negative and unsupportive messages to others, affecting subsequent engagement and feelings of group support | Phone communication with participant and social media decorum rules reminder email sent to all participants. |
| Three participants complained about the frequency of reminders to tweet texts; acceptability data post-intervention showed a preference for fewer text reminders. | “Text fatigue” addressed by decreasing the reminders to tweet from every 24 h of not tweeting to every 72 h of not tweeting part-way through the study. |

Table 4
Participant issues (sent via email).

| Issue category | Number of issues | Number of participants reporting issues |
|----------------|-----------------|----------------------------------------|
| Fitbit device-related | 60 | 25 |
| Syncing trouble | 25 | 15 |
| Account problems | 7 | 6 |
| Skin irritation | 3 | 1 |
| Difficulty getting it on | 2 | 2 |
| General Issues | 19 | 10 |
| Self-resolved Fitbit issue (e.g. “nevermind”) | 5 | 5 |
| Events affecting Fitbit wear or physical activity | 52 | 24 |
| ER or surgeries | 8 | 6 |
| Other health issues | 6 | 4 |
| Other (e.g. lost charger, forgot it) | 38 | 21 |
| Twitter-related | 9 | 8 |
| Account problems | 5 | 5 |
| Number of tweets | 2 | 2 |
| Complaints | 2 | 2 |
Fitbit-only control group’s decrease stands in contrast to a randomized the baseline steps (e.g. ~2800 baseline steps increasing by ~820 is 820.8 (461,3), with a decrease of cooperative goal would be one possibility shown to improve engage would especially effective mid-study period. A group competition or increased total step count for our intervention group was a mean (SE) of increased adherence in real-time (Cadmus-Bertram et al., 2015a). This study also demonstrated feasibility for using consumer-grade devices for daily, automatic physical activity collection over the dur-

The most meaningful effects of exercise are from no drop-off in the tweeting engagement, a boost to increase engagement would be especially effective mid-study period. A group competition or cooperative goal would be one possibility shown to improve engagement in other work (Zhang et al., 2016). This study also demonstrated feasibility for using consumer-grade devices for daily, automatic physical activity collection over the duration of the study period (Wang et al., 2015). The secondary benefits of this data over surveys and accelerometers include: minimized recall bias; increased wear periods; ability to monitor for data quality in real time, rather than waiting for device return (leading to potential data loss incurred in the current study); and ability to monitor intervention adherence in real-time (Cadmus-Bertram et al., 2015a).

Two studies are comparable to our pilot. First, our findings for increased total step count for our intervention group was a mean (SE) of 820.8 (461.3), with a decrease of ~570.9 (430.5) for the control Fitbit only group. While this change in steps is notably small, it is worth noting the baseline steps (e.g. ~2800 baseline steps increasing by ~820 is ~29% increase). The most meaningful effects of exercise are from no activity to any activity (Joseph et al., 2019; Sattelmair et al., 2011). Our Fitbit-only control group’s decrease stands in contrast to a randomized controlled trial (RCT) showing Fitbit only resulted in an increase of 789 (SD = 1979) steps compared with a pedometer (in a similarly aged population after 16 weeks (Cadmus-Bertram et al., 2015b)). Notably, that RCT targeted moderate-to-vigorous physical activity, while ours targeted decreasing prolonged sitting. It is possible that a self-monitoring device such as Fitbit may need to be specifically paired with an exercise intervention targeting behavior to have an impact. The Tweet4Wellness platform could allow for a future study design to test this, comparing motivational differences between decreasing sedentary behavior throughout the day to a single bout of exercise when paired with the Fitbit.

A second comparable study provided within-group challenges and incremental physical activity goal tweets for a 2-month single-arm pilot to private twitter groups of 5–7 undergraduate participants with study-provided Fitbits (Chung et al., 2016). Our study showed similar Fitbit wearing engagement (current study: 95% of the 154 days prescribed, Chung et al.: 99% for 60 days), and number of tweets sent (current study: 0.67 tweets/participant/day, Chung et al.: 0.72 tweets/partici-

Fig. 1. Study Timeline.

Table 5

| Outcome                                | Group                  | Pre Control n = 21 | EOT Control n = 20 | Follow-up Control n = 21 | p-value (Pre-EOT) | p-value (Pre-Follow-up) |
|----------------------------------------|------------------------|--------------------|--------------------|--------------------------|-------------------|-------------------------|
| Total sedentary time (minutes)         | Control                | 988.8 (835.2, 1142.4) | 979.7 (826.8, 1132.6) | 940.1 (786.8, 1093.4) | p = .84           | p = .74                 |
|                                        | Treatment              | 1014.1 (888.0, 1140.1) | 1011.6 (885.8, 1137.5) | 991.0 (862.7, 1119.4) |                  |                         |
| Number of active hours                 | Control                | 4.5 (3.1, 5.9)     | 3.7 (2.3, 5.1)     | 3.9 (2.6, 5.3)   | p = .018           | p = .057                |
|                                        | Treatment              | 3.3 (2.1, 4.4)     | 3.5 (2.3, 4.6)     | 3.5 (2.3, 4.7)   | ES = 0.708          | ES = 0.558              |
| Max sedentary bout length (minutes)    | Control                | 60.2 (51.5, 68.8)  | 61.7 (53.2, 70.3)  | 60.4 (51.8, 69.0) | p = .03            | p = .90                 |
|                                        | Treatment              | 67.5 (60.2, 74.8)  | 69.8 (62.6, 77.0)  | 69.4 (61.8, 77.0) |                  |                         |
| Weighted median sedentary bout length  | Control                | 22.0 (15.7, 28.3)  | 24.6 (18.4, 30.9)  | 23.0 (16.7, 29.2) | p = .33            | p = .62                 |
| (minutes)                              | Treatment              | 27.3 (22.1, 32.6)  | 28.1 (22.9, 33.3)  | 27.2 (21.8, 32.6) |                  |                         |
| Total daily steps                      | Control                | 4735.7 (1848.9, 7622.5) | 4164.82 (1290.8, 7038.8) | 4624.9 (1744.1, 7505.7) | p = .028          | p = .081                |
|                                        | Treatment              | 2819.0 (449.8, 5188.3) | 3639.8 (1273.8, 6005.9) | 3408.52 (995.75, 5821.3) |                  |                         |
| Intensity minutes                      | Control                | 5.4 (–17.5, 28.4)  | 3.4 (–19.4, 26.2)  | 4.4 (–18.4, 27.4) | p = .23            | p = .26                 |
|                                        | Treatment              | 4.5 (–14.4, 23.5)  | 10.0 (–8.9, 28.9)  | 2.7 (–16.7, 22.2) |                  |                         |
| Sitting-to-moving ratio                | Control                | 8.5 (4.2, 12.7)    | 10.6 (6.4, 14.7)   | 8.7 (4.5, 12.9)   | p = .01            | p = .047                |
|                                        | Treatment              | 12.1 (8.5, 15.7)   | 10.1 (6.6, 13.6)   | 10.1 (6.3, 13.8)  | ES = –1.123        | ES = –0.622             |

Control: Fitbit-only.
Treatment: Tweet4Wellness + Fitbit.
p-values are from two-sided Wald tests with \( \alpha = 0.05 \).

\( n \)'s vary based on Fitbit data syncing during time period.

Hedges (Hedges, 2007) effect sizes listed for significant p-values.
Fig. 2. CONSORT diagram.

Fig. 3. Self-reported perceived change over treatment period by condition.
on sedentary behavior questions rather than physical activity questions.

Our sample population was largely Caucasian, which limits generalizability to other ethnic groups. Our sample was quite diverse in age, however, ranging from 36 to 87 years old, averaging 60 years. Pew research noted that 24% of Twitter users are over 50 years old, and only 7% are over 65 years old (Social Media Update, 2013). Most of the participants shared that they were new to Twitter when setting up the accounts for the study. It is possible that our sample had lower engagement than other population subgroups that have more familiarity with Twitter. The novelty of the platform, however, may also be a strength for increasing engagement. Further, while the broad age range (36–87 years old) is inclusive and broad, group rapport and connectedness may have been improved if the group shared similar life phase constraints (e.g., working vs retired).

While we did pilot test the intervention messages to ensure relatability and perceived utility, we did not consult with our participant population on the design of our intervention, a community-based participatory approach strategy. An important step for the next study would be to seek advice from members of the target population potentially obviate implementation obstacles and increase the impact of the intervention.

The unfortunate device failure at baseline removed the possibility of participant-blinded baseline activity data. Our baseline data from Fitbit likely showed higher than usual activity for the participants due to the novelty of the self-tracking, real-time device. This potentially higher-than-normal baseline activity of all participants may have diluted the actual change in behavior, and was conservative to our hypotheses. If consumer-based devices (rather than accelerometers which do not provide data to the participants) are also collecting primary outcomes, future studies should increase baseline data collection to 2 weeks and provide participant education about the importance of accurate baseline data collection.

Another limitation is that we had a large number of outcomes and comparisons. This was intentional however, as this was designed as a pilot study that would inform a larger study with a refined intervention, currently under design. Many of our comparisons were made to provide insight into how our intervention did or did not affect change. Additionally, there is no single measure of sedentary behavior that captures all of the characteristics of max duration, total time sedentary, distribution of activity, and average time spent in prolonged sitting. Further, to address the multiple outcomes and comparisons, we included post-intervention and follow-up timepoints in our models, but only tested changes from baseline to post-intervention, limiting the number of tests for each outcome. As such, we provide careful interpretation so that the reader has the relevant context on which endpoints the study was primarily targeted when addressing the preliminary efficacy of the intervention, and which were providing additional insight into other aspects of success or how our intervention worked (e.g. the effects of engagement on outcomes).

Finally, a risk with any consumer device, the Fitbit company introduced a new element to the device after our study began. “Reminders to move” (vibrating of the device at 10 min before the hour if 250 steps were not achieved) and active hour tracking increased the strength of the control intervention from self-monitoring alone to one with just-in-time behavioral nudges, and specifically around our target behavior. That Tweet4Wellness still had an effect on other outcomes is a testament to the potential of this intervention. Despite the device change, we prioritized keeping Fitbit as an intact treatment, rather than manually overriding features for each participant, to preserve ecological validity and provide a conservative test of our intervention.

Table 6b
Tweet4Wellness Acceptability Survey Data: “what would you change” and planned changes.

| Participant quotes | Modification for next study |
|--------------------|-----------------------------|
| • Make it clear Tweets should only be about project related subjects | Increase education around Twitter and social media etiquette. |
| • Consider issuing guidelines as to what is appropriate to tweet | |
| • When some of the participants posted negative comments it was hard to stay motivated | |
| • Have real people with real exercise issues participate. I thought that people who went scuba diving and walked 14 k did not belong in the group. I was a low step exerciser. | |
| • I would have preferred to get the daily behavior change prompt in a text every morning and not twitter. | |
| • Also I am trying to make a conscious decision to lessen my social media time so this conflicted with that. | |
| • I did not feel alone in this project. It was enlightening to hear the comments of other participants. | |

Table 6a
Tweet4Wellness Acceptability Survey Data: “what did you like?”

- The daily behavior change prompt
- Sharing my own experiences and hearing others felt that I wasn’t alone on the journey. Relatedness and encouragement. Feeling more accountable.
- I realized everyone suffers from the same issues about getting more activity in every day. Since I am retired I was able to really focus on my goal everyday.
- I did not feel alone in this project. It was enlightening to hear the comments of other participants.

Fig. 4. Proportion of participants who synced their Fitbit over entire study period. Each dot represents one day. Line represents Lowess curve.

Fig. 5. Proportion of Treatment participants who tweeted on a given day over treatment period. Each dot represents one day. Line represents Lowess curve.
4.2. Lessons learned

An important objective of this pilot study was to inform refinement of the intervention for a larger trial and recommendations for other similar interventions. One important modification will be to only use Fitbits for data collection devices to avoid costly accelerometer device and data loss and failures. While the outputs available on the Fitbit device are not the raw, triaxial accelerometer gold standards for movement assessment, the devices are sensitive to change within an individual. Additionally, the individual is self-monitoring their own activity based on the Fitbit’s output, not gold-standard raw data from accelerometers; therefore using the Fitbit to also measure change in similar interventions. One important modification will be to only use the intervention for a larger trial and recommendations for other

A second refinement will be to limit the sedentary behavior characteristic outcomes to the following: total steps as a marker of overall movement; number of active hours as a marker of distribution of accelerometers; therefore using the Fitbit to also measure change in similar interventions. One important modification will be to only use the intervention for a larger trial and recommendations for other

To boost the effect of the intervention and increase engagement, one idea is to add a challenge midway through the intervention period. Alternatively, screening for a population that is more comfortable and interested in using social media would likely increase engagement and interaction. Finally, to increase the actual use of the intervention-delivered daily messages, participants could be asked to “like” or “thumbs up” the strategies they try that work for them.

The effects from this intervention were small, and, except for the proportion of sitting:movement, did not hold at the 22-week follow-up. One possible way to increase the duration of the changes and increase the size of the effects is to deliver the evidence-based strategies in a way that is actually learned and used by the participants. The current study simply texted them daily to the group with an engagement prompt. While we do not systematically describe the group tweets here, only 8% of the group tweets actually referred to the evidence-based strategies or answered the prompts to use the strategies, and these were only in the first 2 weeks of the treatment phase. The remaining posts were group support or non-strategy referencing interactions. To improve learning from or using the strategies themselves, future studies should consider evidence-based teaching methods to deliver these behavioral strategies. For example, getting participants to practice the same strategy multiple times throughout the study period will lead to better retention of this strategy later (Benjamin and Tullis, 2010). In addition, future work should include measures such as self-efficacy of behavior change, use of strategies, and other mechanisms of change to identify further places for intervention refinement.

Tweet4Wellness is a feasible, acceptable, and potentially effective way to promote reduction of sedentary behavior via the novel integration of a social media platform and a smart step-tracking device. This first maiden voyage of the intervention provided insight and opportunities for growth to help both the next iteration of the treatment and other similar mHealth interventions in the field.

Funding

The initial study design funding was obtained from the Women’s Heart Health Clinic at Stanford and the Stanford Clayman Institute. Funding to run this pilot study was received from the National Institutes of Health’s National Heart, Lung, and Blood Institute under Award Number K01HL136702. The sponsors were not involved in any of the research or analyses.

Trial registration

ClinicalTrials.gov NCT02958189, https://clinicaltrials.gov/ct2/show/NCT02958189

International Registered Report Identifier (IRRID): DERR1-10.2196/2092.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

An, L., Schillo, B., Soul, J., et al., 2008. Utilization of smoking cessation informational, interactive, and online community resources as predictors of abstinence: cohort study. J. Med. Internet Res. 10 (5), e55 https://doi.org/10.2196/jmir.1018.
Bandura, A., 2001. Social cognitive theory: an agentic perspective. Annu. Rev. Psychol. 52 (1), 1–26.
Benjamin, A.S., Tullis, J., 2010. What makes distributed practice effective? Cogn. Psychol. 61 (3), 228–247. https://doi.org/10.1016/j.cogpsych.2010.05.004.
Boerema, S., van, Velsen L., Hermens, H., 2019. An intervention study to assess potential effect and user experience of an mHealth intervention to reduce sedentary behaviour among older office workers. BMJ Health Care Inform. 26 (1), e100014. https://doi.org/10.1136/bmjhci-2019-100014.
Buckingham, S.A., Williams, A.J., Morrissey, K., Price, L., Harrison, J., 2019. Mobile health interventions to promote physical activity and reduce sedentary behaviour in the workplace: a systematic review. Digit. Health. 5, 2055201601983983. https://doi.org/10.1177/2055201601983983.
Byrom B, Stratton G, Carthy MM, Muehnsen W. Objective measurement of sedentary behaviour using accelerometers. Int. J. Obes. 2016;40(11):1809-1812. doi:10.1038/ijob.2016.136.
Cadmus-Bertram I., Marcus BH, Patterson RE, Parker BA, Morey BL. Use of the Fitbit to measure adherence to a physical activity intervention among overweight or obese, postmenopausal women: self-monitoring trajectory during 16 weeks. JMIR mHealth uHealth. 2015a;3(4):e96. doi:10.2196/mhealth.4229.
Cadmus-Bertram L., Marcus, B.H., Patterson, R.E., Parker, B.A., Morey, B.L. 2015b. Randomized trial of a Fitbit-based physical activity intervention for women. Am. J. Prev. Med. 49 (3), 414–418. http://dx.doi.org/10.1016/j.amepre.2015.01.020.
CDC, 2020. Heat disease facts | cdc.gov. Centers for Disease Control and Prevention. Published September 8. Accessed September 17, 2020. https://www.cdc.gov/heatdisease/facts.htm.
Chung, A.E., Skinner, A.C., Hastie, S.E., errin, E.M., 2016. Tweeting to health: a novel mHealth intervention using Fitbits and twitter to foster healthy lifestyles. Clin. Pediatr. https://doi.org/10.1177/0009922816653385. Published online June 16.
Cole-Lewis, H., Ezeanochie, N., Turgiss, J., 2019. Understanding health behavior technology engagement: pathway to measuring digital behavior change effects. J. Med. Internet Res. 10 (5), e55 https://doi.org/10.2196/jmir.1018.
Davis, R., Campbell, R., Hildon, Z., Hobbs, L., Mitchie, S., 2015. Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. Health Psychol. Rev. 9 (3), 323–344. https://doi.org/10.1080/17437199.2014.941722.
Diaz, K.M., Howard, V.J., Hutto, B., et al., 2017. Patterns of sedentary behavior and mortality in U.S. middle-aged and older adults: a National Cohort Study. Ann. Intern. Med. 167 (7), 465–475. https://doi.org/10.7326/M17-0212.
Dweck, C.S., 1996. Implicit theories as organizers of goals and behavior. In: The Psychology of Action: Linking Cognition and Motivation to Behavior. The Guilford Press, pp. 69–90.
Evenson, K.R., Butler, E.N., Rosamond, W.D., 2014. United States prevalence of physical activity and sedentary behavior among adults with cardiovascular disease. J. Cardiopulm. Rehabil. Prev. 34 (6), 406–419. https://doi.org/10.1097/HCR.0000000000000064.
Fiedler, J., Eckert, T., Wunsch, K., Woll, A., 2020. Key facets to build up eHealth and mHealth interventions to enhance physical activity, sedentary behavior and nutrition in healthy subjects - an umbrella review. BMC Public Health 20 (1), 1605. https://doi.org/10.1186/s12889-020-09765-7.
Giurgiu M, Koch ED, Plotnikoff RC, Ebner-Priemer UW, Reichert M. Breaking up sedentary behavior optimally to enhance mood. [miscellaneous article]. Med. Sci. Sports Exerc.. 2020;52(2):457–465. doi:10.1249/MSS.0000000000002152.
Goldberg, J.H., Kiernan, M., 2005. Innovative techniques to address retention in a behavior-weight-loss trial. Health Educ. Res. 20 (4), 439–447. https://doi.org/10.1093/her/cyq139.
Gollwitzer, P.M., 1999. Implementation intentions: strong effects of simple plans. Am. Psychol. 54 (7), 493.
Hedges, L.V., 2007. Effect sizes in cluster-randomized designs. J. Educ. Behav. Stat. 32 (4), 341–370. https://doi.org/10.3102/1076998606298943.
Joseph, G., Marott, J.L., Torp-Pedersen, C., et al., 2019. Dose-response association between level of physical activity and mortality in normal, elevated, and high blood pressure. Hypertension 74 (6), 1307–1315. https://doi.org/10.1161/HYPERTENSIONAHA.119.13796.
King, A.C., Hekler, E.B., Grieco, L.A., et al., 2013. Harnessing different motivational frames via Mobile phones to promote daily physical activity and reduce sedentary behavior in aging adults. PloS One 8 (4), e62613. https://doi.org/10.1371/journal. pone.0062613.
