

Effectiveness of activity trackers with and without incentives to increase physical activity (TRIPPA): a randomised controlled trial

Eric A Finkelstein, Benjamin A Haaland, Marcel Bilger, Aarti Sahasranaman, Robert A Sloan, Ei Ei Khaing Nang, Kelly R Evenson

Summary

Background Despite the increasing popularity of activity trackers, little evidence exists that they can improve health outcomes. We aimed to investigate whether use of activity trackers, alone or in combination with cash incentives or charitable donations, lead to increases in physical activity and improvements in health outcomes.

Methods In this randomised controlled trial, employees from 13 organisations in Singapore were randomly assigned (1:1:1:1) with a computer generated assignment schedule to control (no tracker or incentives), Fitbit Zip activity tracker, tracker plus charity incentives, or tracker plus cash incentives. Participants had to be English speaking, full-time employees, aged 21–65 years, able to walk at least ten steps continuously, and non-pregnant. Incentives were tied to weekly steps, and the primary outcome, moderate-to-vigorous physical activity (MVPA) bout min per week, was measured via a sealed accelerometer and assessed on an intention-to-treat basis at 6 months (end of intervention) and 12 months (after a 6 month post-intervention follow-up period). Other outcome measures included steps, participants meeting 70 000 steps per week target, and health-related outcomes including weight, blood pressure, and quality-of-life measures. This trial is registered at ClinicalTrials.gov, number NCT01855776.

Findings Between June 13, 2013, and Aug 15, 2014, 800 participants were recruited and randomly assigned to the control (n=201), Fitbit (n=203), charity (n=199), and cash (n=197) groups. At 6 months, compared with control, the cash group logged an additional 29 MVPA bout min per week (95% CI 10–47; p=0·0024) and the charity group an additional 21 MVPA bout min per week (2–39; p=0·0310); the difference between Fitbit only and control was not significant (16 MVPA bout min per week [–2 to 35; p=0·0854]). Increases in MVPA bout min per week in the cash and charity groups were not significantly greater than that of the Fitbit group. At 12 months, the Fitbit group logged an additional 37 MVPA bout min per week (19–56; p=0·0001) and the charity group an additional 32 MVPA bout min per week (12–51; p=0·0013) compared with control; the difference between cash and control was not significant (15 MVPA bout min per week [–5 to 34; p=0·1363]). A decrease in physical activity of –23 MVPA bout min per week (95% CI –42 to –4; p=0·0184) was seen when comparing the cash group with the Fitbit group. There were no improvements in any health outcomes (weight, blood pressure, etc) at either assessment.

Interpretation The cash incentive was most effective at increasing MVPA bout min per week at 6 months, but this effect was not sustained 6 months after the incentives were discontinued. At 12 months, the activity tracker with or without charity incentives were effective at stemming the reduction in MVPA bout min per week seen in the control group, but we identified no evidence of improvements in health outcomes, either with or without incentives, calling into question the value of these devices for health promotion. Although other incentive strategies might generate greater increases in step activity and improvements in health outcomes, incentives would probably need to be in place long term to avoid any potential decrease in physical activity resulting from discontinuation.

Funding Ministry of Health, Singapore.

Introduction

Non-communicable diseases are fast emerging as the major health challenge for the 21st century and are responsible for nearly two-thirds of global deaths annually.¹ Physical inactivity is an important risk factor for non-communicable diseases and has been identified as the fourth leading risk factor for global mortality, contributing to 9% of deaths annually.² Inactivity also imposes substantial costs to governments, insurers, and employers through increased medical expenditures for treating non-communicable diseases and reductions in productivity.³

Wearable activity trackers have greatly enhanced our

ability to track and potentially increase physical activity levels. The affordability, ease of wear, and aesthetics of these devices have increased their popularity. One in ten US adults now owns an activity tracker.⁴ These devices are also increasingly being used by employers, insurers, and governments to promote physical activity.^{5,6}

On the basis of their systematic review, Bravata and colleagues⁷ suggest that pedometers increase physical activity by an average of 26·9% over baseline and generate slight improvements in BMI and systolic blood pressure. However, in most of the studies reviewed, the step differences were measured with a pedometer, which does

Lancet Diabetes Endocrinol 2016

Published Online

October 4, 2016

[http://dx.doi.org/10.1016/](http://dx.doi.org/10.1016/S2213-8587(16)30284-4)

[S2213-8587\(16\)30284-4](http://dx.doi.org/10.1016/S2213-8587(16)30284-4)

See Online/Comment

[http://dx.doi.org/10.1016/](http://dx.doi.org/10.1016/S2213-8587(16)30264-9)

[S2213-8587\(16\)30264-9](http://dx.doi.org/10.1016/S2213-8587(16)30264-9)

Program in Health Services and Systems Research

(Prof E A Finkelstein PhD,

M Bilger PhD,

A Sahasranaman PhD) and

Centre for Quantitative

Medicine (B A Haaland PhD),

Duke-NUS Medical School,

Singapore, Singapore; H Milton

Stewart School of Industrial &

Systems Engineering, Georgia

Institute of Technology,

Atlanta, GA, USA (B A Haaland);

The Graduate School of Medical

and Dental Sciences,

Kagoshima University,

Kagoshima, Japan

(R A Sloan PhD); Lee Kong Chian

School of Medicine, Nanyang

Technological University,

Singapore, Singapore

(E E K Nang PhD);

Gillings School of Global Public

Health, Department of

Epidemiology, University of

North Carolina, Chapel Hill, NC,

USA (Prof K R Evenson PhD);

and Duke Global Health

Institute, Duke University,

Durham, NC, USA

(Prof E A Finkelstein)

Correspondence to:

Prof Eric A Finkelstein, Program

in Health Services and Systems

Research, Duke-NUS Medical

School, Singapore 169857,

Singapore

[eric.finkelstein@duke-nus.edu.](mailto:eric.finkelstein@duke-nus.edu.sg)

sg

Research in context

Evidence before this study

In preparation for this trial, we began with the systematic review of pedometer studies by Bravata and colleagues in 2007. Their review suggests that pedometer use increases physical activity by an average of 26.9% over baseline, and produces slight reductions in BMI and improvements in systolic blood pressure. However, the review raised several concerns that question the veracity of the results. In most of the studies reviewed, the step differences were measured using a pedometer, which does not allow for disentangling differences in wear time from actual increases in steps. Bravata and colleagues further noted that weight loss was not a function of increases in daily steps and that blood pressure reductions were potentially caused by regression to the mean. Moreover, mean intervention duration was only 18 weeks and results were weak for workplace interventions. The authors concluded by stating that large randomised controlled trials are needed to fully elucidate the potential benefits of pedometers. In a 2013 Cochrane review on workplace pedometer interventions, Freak-Poli and colleagues concluded that insufficient evidence was available to assess effectiveness of pedometer-based interventions in the workplace. Market research further reveals that roughly a third of people who purchase activity trackers abandon them within 6 months of purchase. We searched MEDLINE and PreMEDLINE, Econlit, CINAHL, Embase, PsycINFO, Scopus, and the Cochrane Central Register of Controlled Trials for articles published in English between Jan 1, 2007, and April 30, 2013, that used high quality research designs to test the use of incentives to influence behaviours related to physical activity. The search terms used were combinations of the keywords “behavioural economics”, “behavioural interventions”, “incentives”, “financial incentives”,

not allow differences in wear time to be disentangled from actual increases in activity. Weight loss was not a function of increases in daily steps and blood pressure reductions were potentially caused by regression to the mean. The mean intervention duration was only 18 weeks and results were weakest for people who were employed full-time compared with people who were not. In a 2013 Cochrane review⁸ on workplace pedometer interventions, Freak-Poli and colleagues concluded that insufficient data exist to determine the effectiveness of pedometer-based interventions in the workplace. Moreover, market research data reveal that more than 50% of US consumers who purchased activity trackers no longer wear them, with a third abandoning them within 6 months of purchase.⁴ These results call into question the ability of trackers to improve population health.

As originally designed, activity trackers provided timely feedback to wearers on step activity. For many, this feedback had little value once the novelty of the device wore off. People who were consistently active or inactive quickly learned what their typical daily step count was. Consequently, the devices provided little by way of new

and “non-communicable diseases”, “chronic diseases”, “physical activity”.

Added value of this study

To our knowledge, this is the first large-scale, randomised controlled trial that tests the effectiveness of activity trackers in full-time workers, either alone or in combination with one of two incentive strategies (namely, cash and charitable donations) to increase physical activity and improve health outcomes. Economic theory suggests that incentives would extend the period that participants wear activity trackers and boost increases in activity as individuals strive to obtain the rewards. We tested these hypotheses, in total and separately, for insufficiently active and active participants during a 6 month period when incentives are in effect and during the subsequent 6 months without the incentives to test for habit formation. Despite some evidence of the ability of the tracker to improve step activity, we identified no evidence of improvements in health outcomes, either with or without incentives.

Implications of all the available evidence

The preponderance of evidence suggests that activity trackers have little ability to improve health outcomes related to inactivity. Our findings suggest that coupling the tracker with incentives might hold some promise, but that if the incentives are provided in cash they would probably need to be in place long term to avoid an undermining effect resulting from their removal. Future research should investigate alternate incentive strategies that provide a stronger link between the target of the incentives and health outcomes, such as incentivisation of bouts of moderate-to-vigorous physical activity.

information and so it was not surprising that many abandoned them after a few weeks or months. Manufacturers are aware of this limitation and now couple the devices with social media and other sources of information and feedback to increase their value. As such, consumers might be less likely to abandon the new wearables, which should increase their effectiveness.

Another strategy that could increase the value of activity trackers is to couple them with incentives for reaching activity goals. Economic theory predicts that incentives, by increasing the benefits of physical activity, would generate increased physical activity as individuals strive to obtain set rewards. Incentives need not involve individual rewards. An alternative approach could be to earmark incentives to a designated charity, as is commonly done in walkathons.⁹ Although we could not find any published reports of studies testing this approach, it might promote sustained behaviour change while not violating social norms (ie, creating a financial transaction in response to activities for which this is not typically done),^{10,11} which might occur with the uncommon practice of paying people to exercise. Charitable incentives tied to health behaviours are also

consistent with the theory of reasoned action, which states that people will perform a behaviour if they evaluate it positively and believe it important that others think they should perform it.¹⁰ Charitable incentives might also be preferred by employers in view of the positive public relations and tax deductions for donating to charity.

The goal of this randomised controlled trial (trial of economic incentives to promote physical activity [TRIPPA]) is to test the extent to which an activity tracker, with or without modest incentives paid either in cash or via charitable donations, can increase physical activity and improve health outcomes among working people during a 6 month period, and to quantify the extent to which any improvements are sustained during the subsequent 6 months after incentives are removed (termed habit formation). We hypothesised that the activity tracker would have a slight positive effect on physical activity and health outcomes compared with the control group (no tracker or incentives) at 6 months, but that greater effectiveness would result from the incentives, with cash incentives hypothesised to have the largest effects. The value of cash and charity incentives would be strengthened if results showed evidence of habit formation. Although not expressly prespecified in the study protocol, we tested these hypotheses in total and separately for active and insufficiently active participants. This is an important stratification because providing incentives for participants who are already active might not be viewed as a good use of resources by funders.

Methods

Study design and participants

TRIPPA was a four-arm, 6 month randomised controlled trial with a 6 month post-intervention follow-up period, done in 13 organisations spanning many industries and sectors of government, to investigate the effect of an activity tracker, with or without cash or charitable incentives, on physical activity and health outcomes among full-time workers in Singapore. Although these organisations spanned diverse industries, most employees at these organisations held desk jobs that provided little opportunity for work-related physical activity. The study design has been described in detail elsewhere.⁹ Despite being a highly walkable city, 42% of full-time workers surveyed from Singapore's general population were insufficiently active based on self-reported physical activity.¹²

To be eligible, participants had to be English speaking (English is one of four official languages of Singapore), full-time workers, aged 21–65 years, willing to be randomly allocated to one of the four study groups, willing to wear an activity tracker for the duration of the study, able to walk at least ten steps continuously, and non-pregnant. Organisations were engaged through existing contacts and “cold calls”. If companies responded positively, a study briefing was done to apprise the management of study details. Once we received confirmation of participation

from the management team, recruitment materials (eg, electronic direct mails, posters, and newsletters) communicating information about the study were disseminated to employees through internal channels unique to each organisation. Participants who answered yes to any of the questions on the Physical Activity Readiness Questionnaire,¹³ or had a BMI of more than 40 kg/m² were required to obtain physician consent before enrolment. To dissuade marginally interested participants from joining, participants were required to pay an enrolment fee of S\$10.00 (US\$1=\$\$1.34) upon giving informed consent by signing the consent form. We had originally proposed to collect a refundable deposit of S\$20.00 for the Fitbit Zip in addition to the enrolment fee at enrolment; participants subsequently randomly assigned to the control group would have been immediately refunded the S\$20.00. However, this requirement was waived before the start of recruitment after consultations with worksite representatives revealed that a total fee of S\$30.00 might be prohibitive for some employees.

The study protocol (appendix) was approved by the institutional review board at the National University of Singapore.

See Online for appendix

Randomisation and masking

We randomly assigned participants to one of four study groups: control (no tracker or incentives), activity tracker and website (Fitbit), tracker plus charitable incentives (charity), and tracker plus cash incentives (cash). Although randomisation was done at the individual level, we allowed exercise companions to sign up in groups of up to four members and be randomly allocated to the same study group. Randomisation of groups of a particular size was done in blocks of size four to ensure approximate balance between study groups for participants of each group size. We randomly assigned groups (with 1–4 members) to one of the four study groups in a 1:1:1:1 ratio, using a computer generated assignment schedule prepared by the statistician. Randomisation envelopes containing a slip of paper indicating the study group were prepared by research staff not involved in random allocation. The envelopes were arranged in sequential order for each stratum (ie, groups of 1–4 members), and the top-most envelope was picked by the study coordinator on the basis of the stratum to which the participant belonged. Because of the nature of the intervention, the participants and study coordinator could not be blinded to the study group assignment. However, team members involved in assessing outcomes in participants and data analysts were blinded to group assignment.

Procedures

Upon enrolment, all participants received educational booklets, published by the Singapore Health Promotion Board, entitled Active for Life¹⁴ and Brisk Walk Your Way to Better Health.¹⁵ These booklets give readers a better understanding of the benefits of and strategies for

increasing physical activity. Participants in the Fitbit, charity, and cash groups also received the Fitbit Zip wireless activity tracker (Fitbit, San Francisco, CA) and access to the Fitbit website, which provides feedback on daily step activity and additional features (eg, badges, competitions) aimed to sustain increased activity levels. Fitbit Zip is a small, waist-worn pedometer with high validity and reliability.¹⁶ We instructed participants to wear the Fitbit in the same position every day. We told participants in the two incentive groups that they could earn weekly incentives: S\$15 if they logged between 50 000 and 70 000 steps per week or S\$30 if they logged 70 000 or more steps per week.

To prevent injuries, a maximum of 20 000 steps per day were counted. The key difference between the incentive groups was that the cash group participants received the incentives for themselves, whereas the charity group participants raised money for a charity of their choice from a list of 13 local charities representing causes ranging from child welfare to wildlife conservation. Participants in the cash and charity groups also had access to the TRIPPA website where they could track incentives earned.

To keep the other participants engaged in the study, control and Fitbit participants earned a weekly participation payment of S\$4, irrespective of number of steps recorded. We distributed these payments, and the incentive payments, at the participants' worksites every 4–6 weeks. Donations to charities were made via an online portal. Before receiving incentive payments, we asked participants in the charity and cash groups to sign an oath affirming that they (or their designated charity) were being paid for their own physical activity. Research reveals that such oaths reduce the likelihood of participants engaging in dishonest behaviour.¹¹

All physical activity outcomes, including steps, were measured for participants in all groups via sealed accelerometers (ActiGraph triaxial GT-3x+ accelerometer [ActiGraph, Pensacola, FL, USA]) at baseline, 6 months, and 12 months. We instructed participants to wear the accelerometer for at least 10 waking hours on each day for 7 consecutive days during each assessment interval. We quantified mean counts per min measured by the accelerometer by dividing the sum of total counts by the number of min per day of wear time across all adherent days (defined as ≥ 10 h/day of wear). We processed accelerometer data in R (version 3.1.2; R Core Team, Vienna, Austria) using the accelerometry package for triaxial accelerometers to obtain standard variables based on vector magnitude counts.¹⁷ We defined moderate-to-vigorous physical activity (MVPA) as a vector magnitude of 2690 counts per min or higher.¹⁸ We defined MVPA bout min as a total of 10 or more consecutive min at the vector magnitude cutoff points, with allowance for interruptions of 1–2 min at a magnitude less than the cutoff point.¹⁹

To measure health outcomes, we recorded participants' height (Seca 217 Mobile Stadiometer, Seca Deutschland, Hamburg, Germany), weight (Seca 869 Mobile Floor

Scale, Seca GmBH, Hamburg, Germany), blood pressure, and heart rate (Welch-Allyn Spot Vital Signs BP monitor, Welch Allyn, NY, USA) at baseline. We recorded the same measurements, with the exception of height, during the follow-up assessments at 6 months and 12 months.

To encourage participants to attend follow-up assessments, we gave participants in all four groups a S\$25 supermarket voucher and a one in ten chance of receiving a S\$50 voucher for completing the month 6 assessment, and a S\$35 voucher and one in ten chance of receiving a S\$50 voucher for completing the month 12 assessment.

Outcomes

We defined the primary outcome as MVPA bout min per week measured via accelerometry, assessed at 6 months (when the intervention ended) and at 12 months (post-intervention follow-up period). We focused on weekly MVPA bout min as our primary outcome because steps during MVPA bouts are more likely than less vigorous activity to affect health outcomes. However, because incentives were tied to meeting weekly step goals, we included mean daily steps and percentage of participants who met the 70 000 weekly step goal as key secondary outcomes. The health outcomes we measured included weight, systolic blood pressure, cardiorespiratory fitness (maximum oxygen consumption [VO_{2max}]; measured by a non-exercise testing of cardiorespiratory fitness [NET-F]),²⁰ and quality of life (measured by the EuroQoL EQ-5D-5L questionnaire).²¹ We converted the EQ-5D-5L responses to quality-of-life weights using the Thailand EQ-5D-5L crosswalk value sets.²² Other secondary outcomes assessed were weekly steps, sedentary or light physical activity min per week, moderate physical activity min per week, vigorous physical activity min per week, equivalence of moderate physical activity min per week, participants meeting physical activity guideline of 150 or more equivalence of moderate physical activity min per week, and participants meeting physical activity guideline of 10 000 or more steps per day. The results for these secondary outcomes are included in the appendix. Because of the low risks involved, the institutional review board did not require collection and reporting of adverse events.

Statistical analysis

The study is powered to detect a difference of at least 30 MVPA bout min per week between groups because increases less than this magnitude are unlikely to generate meaningful health improvements.²³ On the basis of our pilot data (unpublished), this increase translates into a relative effect size of 0.35 SD difference in mean MVPA bout min per week between groups.²⁴ Assuming 20% attrition at month 12 and an intraclass correlation coefficient of 0.075, which is consistent with intraclass correlation coefficients for similar worksite-based behavioural health interventions,²⁵ we estimated that we would require 80 clusters of mean cluster size

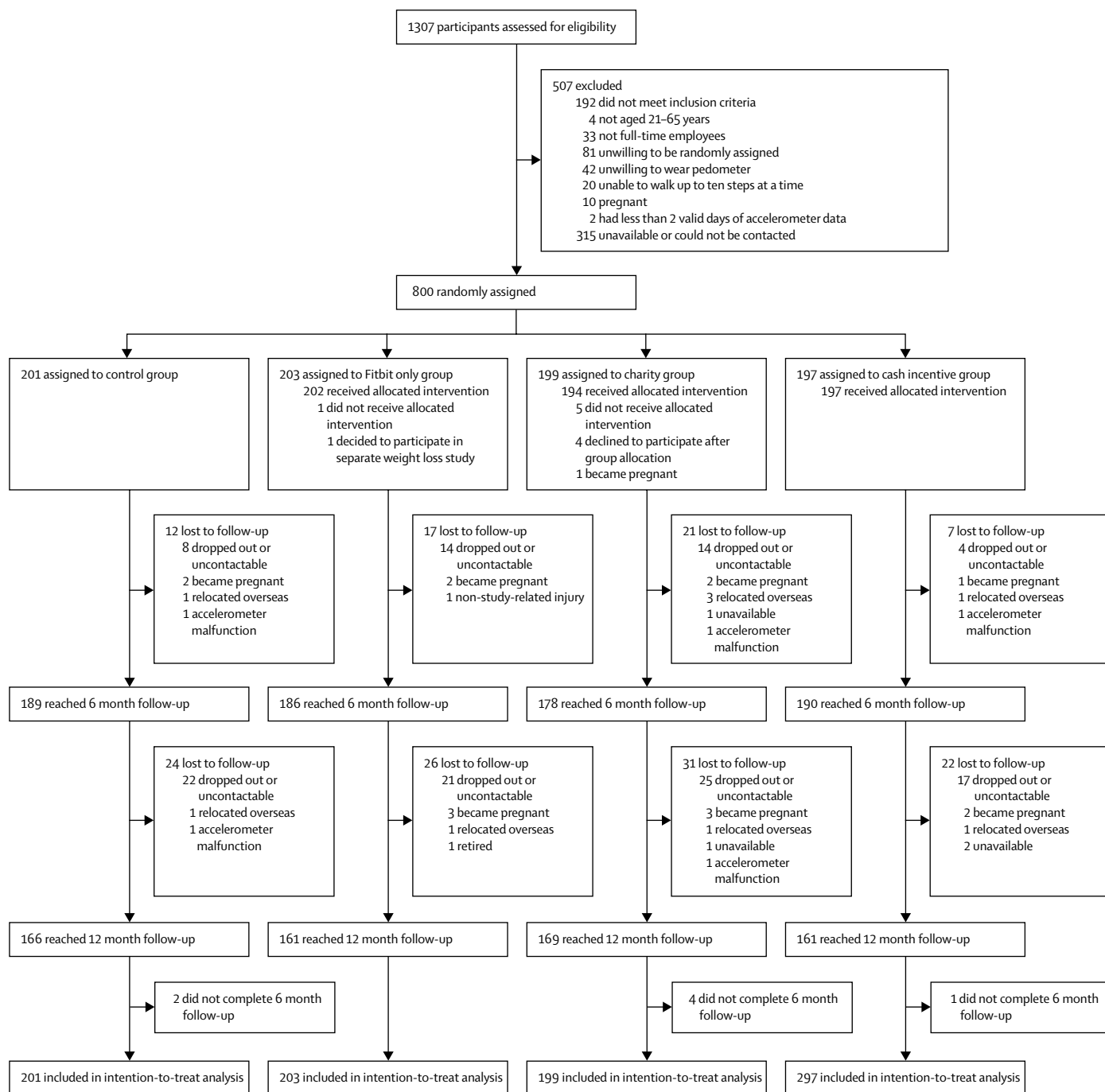


Figure 1: Trial profile

2.5 per group. Mean cluster sizes of less than 2.5 would result in greater power. This translates to 200 participants per group, bringing the total sample size for recruitment to 800.

Analyses were done on an intention-to-treat basis and outcomes were assessed both for the whole study population and for active and insufficiently active

respondents separately. Participants were classified as active or insufficiently active on the basis of their response to the NET-F at baseline. Those who self-reported no physical activity or MVPA up to the 60 min per week category were classified as insufficiently active. Those who reported performing at or above the 1–3 h of MVPA per week category were classified as active.

All statistical analyses were based on estimates pooled²⁶ across 20 multiple imputed datasets²⁷ with each complete dataset based on 50 iterations of predictive mean matching imputation of each dependent variable's missing values conditional on all other variables. Each of the 20 complete

datasets was formed by looping across variables for 50 iterations. Within each iteration, each variable's non-missing values were regressed on all other variables via stepwise variable selection using the Akaike information criterion, then noisy predictions for the variable's missing

	Control (n=201)	Fitbit (n=203)	Charity (n=199)	Cash (n=197)
Age (years)	35.6 (8.6)	35.4 (8.3)	35.5 (8.6)	35.5 (8.4)
Sex				
Male	89 (44%)	103 (51%)	94 (47%)	84 (43%)
Female	112 (56%)	100 (49%)	105 (53%)	113 (57%)
Ethnicity				
Chinese	138 (69%)	133 (66%)	145 (73%)	132 (67%)
Malay	10 (5%)	13 (6%)	5 (3%)	8 (4%)
Indian	30 (15%)	26 (13%)	17 (9%)	32 (16%)
Other	19 (9%)	26 (13%)	27 (14%)	24 (12%)
Not declared	4 (2%)	5 (2%)	5 (3%)	1 (1%)
Education				
High school or lower	12 (6%)	9 (4%)	7 (4%)	11 (6%)
Some college	31 (15%)	27 (13%)	21 (11%)	28 (14%)
College graduate	98 (49%)	100 (49%)	100 (50%)	80 (41%)
Postgraduate	56 (28%)	58 (29%)	60 (30%)	73 (37%)
Other	0	3 (1%)	6 (3%)	4 (2%)
Not declared	4 (2%)	6 (3%)	5 (3%)	1 (1%)
Monthly personal income				
<\$5000	103 (51%)	93 (46%)	92 (46%)	92 (47%)
\$5000–\$9999	41 (20%)	46 (23%)	43 (22%)	37 (19%)
≥\$10 000	12 (6%)	15 (7%)	8 (4%)	15 (8%)
Prefer not to say	40 (20%)	43 (21%)	50 (25%)	52 (26%)
Don't know	1 (<1%)	1 (<1%)	1 (1%)	0
Not declared	4 (2%)	5 (2%)	5 (3%)	1 (1%)
Monthly household income				
<\$5000	52 (26%)	49 (24%)	49 (25%)	53 (27%)
\$5000–\$9999	56 (28%)	57 (28%)	49 (25%)	41 (21%)
≥\$10 000	26 (13%)	29 (14%)	30 (15%)	30 (15%)
Prefer not to say	46 (23%)	49 (24%)	53 (27%)	58 (29%)
Don't know	17 (8%)	14 (7%)	12 (6%)	14 (7%)
Not declared	4 (2%)	5 (2%)	5 (3%)	1 (1%)
Housing				
Private	141 (70%)	157 (77%)	144 (72%)	144 (73%)
Public	54 (27%)	40 (20%)	48 (24%)	48 (24%)
Other	2 (1%)	1 (<1%)	2 (1%)	4 (2%)
Not declared	4 (2%)	5 (2%)	5 (3%)	1 (1%)
Participants who were insufficiently active at baseline*	129 (68%; 61–75)	121 (63%; 56–70)	125 (66%; 58–72)	123 (63%; 56–70)
MVPA bout min per week				
Full sample	121 (105–137)	131 (114–148)	114 (100–127)	131 (115–148)
Insufficiently active	98 (82–115)	99 (79–119)	101 (85–117)	108 (88–127)
Active	159 (128–190)	182 (150–214)	141 (117–166)	174 (145–204)
Daily steps				
Full sample	8030 (7680–8380)	8000 (7670–8330)	7780 (7460–8090)	8300 (7950–8650)
Insufficiently active	7620 (7270–7970)	7550 (7160–7940)	7530 (7150–7900)	7770 (7370–8170)
Active	8670 (7910–9420)	8660 (8040–9270)	8280 (7690–8870)	9240 (8600–9880)

(Table 1 continues on next page)

	Control (n=201)	Fitbit (n=203)	Charity (n=199)	Cash (n=197)
(Continued from previous page)				
Participants meeting $\geq 70\,000$ steps per week				
Full sample	31 (17%; 12–23)	37 (21%; 15–27)	25 (14%; 9–19)	37 (20%; 15–27)
Insufficiently active	16 (12%; 7–19)	17 (15%; 9–23)	15 (12%; 7–19)	15 (14%; 8–21)
Active	15 (28%; 17–41)	20 (28%; 18–40)	10 (15%; 8–26)	22 (32%; 22–45)
Weight (kg)	66.4 (64.4–68.3)	66.9 (64.8–69.0)	66.5 (64.6–68.4)	64.5 (62.6–66.4)
Overweight or obese participants	114 (57%; 50–64)	114 (56%; 49–63)	121 (61%; 54–68)	105 (53%; 46–60)
Systolic blood pressure (mm Hg)	116 (113–118)	114 (111–116)	115 (112–117)	115 (113–118)
Patients with hypertension	19 (9%; 5–14)	13 (6%; 3–10)	17 (9%; 5–13)	20 (10%; 6–14)
NET-F VO_{2max} (mL/kg per min)	35.0 (34.3–35.6)	35.7 (35.0–36.3)	35.3 (34.7–36.0)	35.4 (34.8–36.1)
EQ-5D-5L quality-of-life score	0.89 (0.88–0.91)	0.89 (0.87–0.91)	0.88 (0.86–0.90)	0.89 (0.87–0.91)

Data are mean (SD), n (%), mean (95% CI), and n (%; 95% CI). Dollars are S\$. This table has been partly reproduced from reference 9. All baseline levels of outcomes are reported before multiple imputation. Participants who reported performing up to 60 min of MVPA per week were classified as insufficiently active. Participants who reported performing at or above the 1–3 h of MVPA per week category were classified as active. Insufficiently active and active participants are subsets of the total sample. Participants with BMI ≥ 23 kg/m² are categorised as overweight. Participants with systolic blood pressure of ≥ 140 mm Hg or a diastolic blood pressure of ≥ 90 mm Hg, or both, are classified as hypertensive. VO_{2max} is the maximum rate of oxygen consumption as measured during exercise and reflects the physical fitness of an individual. EQ-5D-5L is a standardised means of measuring generic health status. MVPA=moderate-to-vigorous physical activity. NET-F=non-exercise testing for cardiorespiratory fitness. VO_{2max} =maximum oxygen consumption. *Participants were classified as insufficiently active or active on the basis of their response to the NET-F.

Table 1: Baseline characteristics and levels of primary and secondary outcome measures by study group

components were generated from the regression model, and matched to their nearest non-missing entry.

Point estimates along with clustered standard errors within each multiple imputed dataset were based on generalised linear mixed-effects models with random effects for worksite, group, and, if the outcome was a repeated measure, individual participant. We adjusted comparisons across study groups at each assessment for baseline outcome value, age, sex, race, and for accelerometer wear time-adjusted estimates, baseline 6 month and 12 month wear time. All estimates are produced at the mean value of the independent variables, including mean wear time for the wear time-adjusted estimates.

At month 6, we noticed a significantly higher wear time effect for those in the cash group compared with the other groups (appendix). This effect probably occurred because participants wore the accelerometer and Fitbit together and increased their wear time or physical activity, or both, relative to baseline in an effort to meet the reward threshold. We present results unadjusted for accelerometer wear time as our primary analysis. Although not expressly stated in the study protocol, we also included results adjusted for wear time in the appendix because we cannot disentangle the competing hypotheses of increased wear time or physical activity, or both. We did the statistical analyses in R (version 3.1.0).

This study is registered with ClinicalTrials.gov, number NCT01855776.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Between June 13, 2013, and Aug 15, 2014, we recruited 800 full-time employees from 15 worksites. 201 participants were randomly assigned to control, 203 to the Fitbit group, 199 to the charity group, and 197 to the cash group (figure 1). Even though participants were allowed to enrol in groups of up to four people, 771 (96%) participants enrolled as singletons. The mean age of participants was 35 years and 370 (46%) were men. Consistent with Singapore's population demographics, most participants were of Chinese ethnic origin (n=548 [69%]), followed by Indian ethnic origin (n=105 [13%]), and Malay ethnic origin (n=36 [5%]). Most of the participants (n=625 [78%]) were college graduates or postgraduates and 48% had monthly personal incomes of less than S\$5000 (table 1). At baseline 498 (62%) of all participants were insufficiently active and 270 (34%) were active; 454 (57%) were overweight or obese, and 69 (9%) were hypertensive (table 1).

After 6 months, 57 (7%) participants were lost to follow-up; 12 in the control group, 17 in the Fitbit group, 21 in the charity group, and seven in the cash group (figure 1). After 12 months, 153 (19%) participants were lost to follow-up; 34 in the control group, 43 in the Fitbit group, 48 in the charity group, and 28 in the cash group. Logistic regressions revealed that sex and ethnic origin were the only sociodemographic predictors of participants being lost to follow-up at month 12. Women participants had 1.8 (95% CI 1.2–2.8) times greater odds of dropping out compared with men, and the odds of dropping out were 3.1 (1.9–5.3) times greater in participants of Indian ethnic origin and 1.8 (1.1–3.2) times greater in participants of other ethnic origin compared with participants of Chinese ethnic origin (appendix). All participants were included in the analyses for all primary and secondary outcomes.

Compared with the Fitbit group, participants in the cash and charity groups who wore the Fitbit logged more daily steps during each week of the intervention (figure 2). The mean daily step count among wearers was 11010 steps in the cash group, 9280 in the charity group, and 8550 in the Fitbit group. Even after 1 week, dropout rates were greater in the Fitbit and charity groups (figure 2). By 6 months, only 249 (62%) of 402 participants in the Fitbit or charity

groups were wearing the device, whereas 174 (88%) of 197 participants in the cash group were still wearing the device. However, when incentives were no longer offered, non-wear increased substantially in the cash group; by month 12, only a mean of about 10% of participants in all three groups were still wearing the device.

Between baseline and month 6, participants in the control group showed a reduction in mean daily steps and

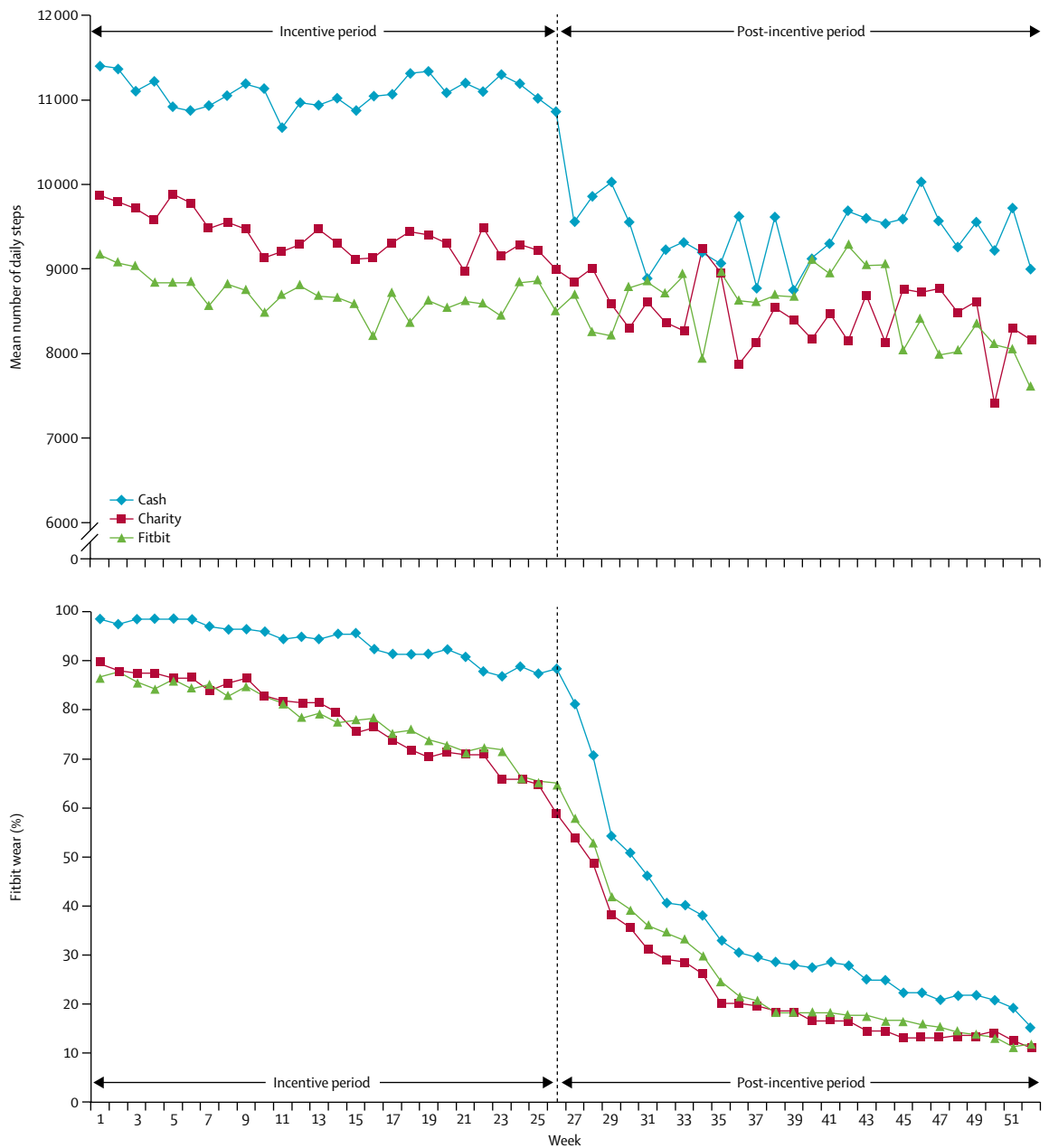


Figure 2: Mean number of daily steps and percentage of Fitbit Zip wear, by study group

Percentage of Fitbit wear indicates the percentage of participants who wore the Fitbit for at least 1 day in that week. Any day during which a participant logged 500 or more steps on the Fitbit was regarded as a valid day and constituted wearing the tracker for that day. The control group is not included in these graphs because control group participants did not receive the Fitbit activity tracker.

	Difference from baseline, by study group				Between-group comparisons					
	Control	Fitbit	Charity	Cash	Fitbit vs control	Charity vs control	Charity vs Fitbit	Cash vs control	Cash vs Fitbit	Cash vs charity
MVPA bout min per week										
Full sample	-16 (-29 to -3; p=0.0228)	0 (-13 to 13; p=0.9626)	5 (-9 to 18; p=0.5068)	13 (-1 to 26; p=0.0643)	16 (-2 to 35; p=0.0854)	21 (2 to 39; p=0.0310)	5 (-13 to 23; p=0.6492)	29 (10 to 47; p=0.0024)	13 (-6 to 31; p=0.1800)	8 (-11 to 26; p=0.3889)
Insufficiently active	-3 (-19 to 13; p=0.7136)	7 (-9 to 23; p=0.3707)	10 (-6 to 26; p=0.2847)	22 (5 to 38; p=0.0096)	10 (-12 to 32; p=0.3439)	12 (-9 to 34; p=0.2758)	2 (-19 to 24; p=0.8937)	24 (3 to 46; p=0.0254)	14 (-8 to 36; p=0.2024)	12 (-10 to 34; p=0.2561)
Active	-39 (-65 to -14; p=0.0033)	-13 (-37 to 11; p=0.3048)	-4 (-29 to 21; p=0.7806)	-3 (-27 to 21; p=0.8186)	27 (-8 to 62; p=0.1513)	36 (0 to 71; p=0.0572)	9 (-26 to 43; p=0.6148)	37 (2 to 71; p=0.0473)	10 (-24 to 44; p=0.5738)	1 (-34 to 36; p=0.9650)
Mean daily steps										
Full sample	-480 (-830 to -120; p=0.0099)	-130 (-490 to 230; p=0.4495)	-300 (-660 to 70; p=0.1138)	570 (210 to 930; p=0.0016)	340 (-100 to 790; p=0.1362)	180 (-270 to 630; p=0.4319)	-160 (-610 to 280; p=0.4943)	1050 (600 to 1490; p<0.0001)	700 (260 to 1150; p=0.0015)	870 (420 to 1320; p=0.0001)
Insufficiently active	-180 (-600 to 240; p=0.4071)	-130 (-550 to 290; p=0.4721)	-180 (-600 to 250; p=0.3863)	460 (40 to 890; p=0.0336)	50 (-470 to 570; p=0.9267)	0 (-510 to 520; p=0.9780)	-50 (-560 to 470; p=0.9063)	640 (120 to 1160; p=0.0152)	590 (70 to 1110; p=0.0218)	640 (110 to 1170; p=0.0145)
Active	-960 (-1610 to -310; p=0.0037)	-130 (-7450 to 490; p=0.7524)	-510 (-1160 to 140; p=0.1496)	700 (80 to 1320; p=0.0227)	830 (-10 to 1670; p=0.0431)	450 (-410 to 1300; p=0.2785)	-380 (-1220 to 450; p=0.3702)	1660 (820 to 2500; p=0.0001)	830 (10 to 1650; p=0.0505)	1210 (380 to 2050; p=0.0052)
Participants meeting ≥70 000 steps per week target										
Full sample	-6% (-11 to -1; p=0.0216)	-1% (-7 to 5; p=0.7028)	3% (-4 to 9; p=0.3836)	8% (0 to 17; p=0.0458)	4% (-4 to 12; p=0.1460)	8% (0 to 17; p=0.0238)	4% (-5 to 13; p=0.3700)	14% (4 to 24; p=0.0025)	10% (-1 to 20; p=0.0985)	6% (-5 to 16; p=0.4984)
Insufficiently active	-2% (-9 to 4; p=0.4555)	1% (-7 to 9; p=0.9133)	0% (-7 to 7; p=0.9168)	8% (-2 to 17; p=0.1111)	3% (-7 to 13; p=0.5387)	2% (-8 to 12; p=0.6464)	-1% (-12 to 10; p=0.8804)	10% (-2 to 22; p=0.1022)	7% (-6 to 19; p=0.2940)	8% (-4 to 20; p=0.2402)
Active	-10% (-19 to -1; p=0.0088)	-4% (-15 to 7; p=0.4921)	9% (-5 to 23; p=0.1559)	10% (-6 to 25; p=0.1997)	6% (-8 to 21; p=0.1176)	19% (3 to 36; p=0.0047)	13% (-5 to 38; p=0.1332)	20% (2 to 38; p=0.0048)	14% (-5 to 33; p=0.1703)	1% (-20 to 21; p=0.7743)

Data are change in accelerometer-derived endpoints (95% CI; p value). Test of endpoints at month 6 were adjusted for baseline endpoint value, age, sex, and ethnic origin. MVPA=moderate-to-vigorous physical activity.

Table 2: Change in steps and MVPA bouts at month 6 (unadjusted for wear time)

MVPA bout min per week (table 2). By contrast, charity and Fitbit participants showed stable step and MVPA bout min per week during this time period. Only participants in the cash group showed increases in physical activity from baseline: mean daily steps increased by 570 steps (95% CI 210–930; p=0.0016) and were statistically greater for both active and insufficiently active participants. Cash participants were also 8% (95% CI 0–17; p=0.0458) more likely at month 6 to meet the 70 000 weekly step target as compared with baseline. However, step increases only translated into a significant increase (22 min [95% CI 5–38; p=0.0096]) in MVPA bout min per week for participants who were insufficiently active at baseline.

The charity group was no more effective at increasing MVPA bout min per week than the Fitbit group at the conclusion of the incentive period at 6 months (table 2). Both groups seemed to experience less deterioration in physical activity compared with control, but the results were only significant for the charity group. After controlling for baseline differences, charity participants engaged in an additional 21 MVPA bout min per week (95% CI 2–39; p=0.0310) compared with control. No significant difference in MVPA bout min per week was found between the Fitbit and control groups.

Even after controlling for baseline differences, both insufficiently active and active participants in the cash

group showed increases in mean daily steps compared with participants in the other groups at the conclusion of the incentive period at 6 months (table 2). However, for MVPA bout min per week, only the differences from control were significant (table 2). For the full sample, cash participants logged an additional 29 MVPA bout min per week (95% CI 10–47; p=0.0024) compared with control at the 6 month assessment point: 24 MVPA bout min per week (3–46; p=0.0254) more for those who were insufficiently active at baseline and 37 MVPA bout min per week (2–71; p=0.0473) more for those who were active.

Results adjusted for accelerometer wear time (appendix) were similar to the results before this adjustment, although the magnitude of the differences and significance were attenuated between the cash and the other groups. During the 6 month incentive period, participants in the charity group earned S\$320.00 for charity on average, with 36% of the total earnings for charity going to participants who were already active at baseline. The participants in the cash group earned S\$620.00 on average, with 62% of the cash earned by those who were already active at baseline.

At 12 months, 6 months after the incentives concluded, physical activity among control participants remained below baseline, with results amplified for MVPA bout

	Difference from baseline by, study group				Between-group comparisons					
	Control (n=201)	Fitbit (n=203)	Charity (n=199)	Cash (n=197)	Fitbit vs control	Charity vs control	Charity vs Fitbit	Cash vs control	Cash vs Fitbit	Cash vs charity
MVPA bout min per week										
Full sample	-22 (-36 to -7; p=0.0032)	16 (2 to 30; p=0.0301)	10 (-4 to 24; p=0.1751)	-7 (-21 to 7; p=0.3403)	37 (19 to 56; p=0.0001)	32 (12 to 51; p=0.0013)	-6 (-25 to 14; p=0.5607)	15 (-5 to 34; p=0.1363)	-23 (-42 to -4; p=0.0184)	-17 (-36 to 2; p=0.0820)
Insufficiently active	-8 (-23 to 7; p=0.2967)	16 (0 to 32; p=0.0547)	17 (2 to 32; p=0.0231)	2 (-13 to 18; p=0.7891)	24 (2 to 46; p=0.0353)	25 (4 to 46; p=0.0179)	2 (-20 to 24; p=0.8802)	10 (-12 to 32; p=0.3679)	-14 (-36 to 9; p=0.2298)	-15 (-37 to 6; p=0.1663)
Active	-51 (-82 to -20; p=0.0012)	17 (-12 to 46; p=0.2575)	-1 (-32 to 29; p=0.9146)	-25 (-54 to 4; p=0.0881)	68 (30 to 106; p=0.0005)	49 (11 to 88; p=0.0127)	-18 (-56 to 20; p=0.3403)	26 (-13 to 64; p=0.1890)	-42 (-79 to -5; p=0.0246)	-24 (-61 to 14; p=0.2160)
Mean daily steps										
Full sample	-480 (-880 to -80; p=0.0198)	-30 (-420 to 370; p=0.8950)	-70 (-480 to 330; p=0.7161)	30 (-370 to 430; p=0.8853)	450 (-10 to 910; p=0.0561)	400 (-60 to 860; p=0.0856)	-50 (-510 to 420; p=0.8398)	500 (50 to 960; p=0.0289)	60 (-400 to 510; p=0.8100)	100 (-350 to 560; p=0.6519)
Insufficiently active	-180 (-610 to 260; p=0.4257)	0 (-450 to 460; p=0.9876)	160 (-280 to 600; p=0.4886)	40 (-400 to 480; p=0.8626)	180 (-370 to 730; p=0.5213)	330 (-190 to 850; p=0.2118)	150 (-400 to 700; p=0.5883)	220 (-310 to 740; p=0.4237)	40 (-530 to 600; p=0.9016)	-120 (-640 to 410; p=0.6658)
Active	-1020 (-1730 to -310; p=0.0048)	-40 (-720 to 640; p=0.9022)	-420 (-1120 to 290; p=0.2453)	-60 (-740 to 620; p=0.8658)	980 (120 to 1840; p=0.0256)	610 (-270 to 1480; p=0.1746)	-370 (-1220 to 470; p=0.3875)	960 (90 to 1840; p=0.0305)	-20 (-860 to 830; p=0.9702)	360 (-480 to 1190; p=0.4009)
Participants meeting ≥70 000 steps per week target										
Full sample	-6% (-11 to -1; p=0.0172)	-1% (-8 to 5; p=0.7658)	4% (-4 to 11; p=0.3129)	4% (-4 to 12; p=0.3611)	5% (-3 to 14; p=0.1031)	10% (1 to 19; p=0.0195)	5% (-6 to 15; p=0.3534)	10% (0 to 19; p=0.0149)	5% (-6 to 15; p=0.3913)	0% (-11 to 11; p=0.8793)
Insufficiently active	-3% (-9 to 3; p=0.3743)	0% (-8 to 8; p=0.9817)	1% (-7 to 8; p=0.8960)	5% (-5 to 15; p=0.2809)	3% (-7 to 13; p=0.5069)	3% (-7 to 13; p=0.4709)	0% (-11 to 11; p=0.9384)	8% (-4 to 20; p=0.1608)	5% (-8 to 18; p=0.4567)	5% (-8 to 17; p=0.5094)
Active	-11% (-21 to -1; p=0.0128)	-3% (-16 to 9; p=0.6315)	11% (-5 to 26; p=0.1591)	1% (-13 to 15; p=0.8848)	8% (-8 to 24; p=0.0873)	22% (3 to 41; p=0.0075)	14% (-6 to 34; p=0.1741)	12% (-5 to 29; p=0.0339)	4% (-14 to 22; p=0.6481)	-10% (-30 to 11; p=0.3218)

Data are change in accelerometer-derived endpoints (95% CI; p value). Test of endpoints at month 12 were adjusted for baseline endpoint value, age, sex, and ethnic origin. MVPA=moderate-to-vigorous physical activity.

Table 3: Change in steps and MVPA bouts at month 12 (unadjusted for wear time)

min per week due to a larger reduction from baseline among active participants whose MVPA bout min per week reduced by 51 min (95% CI -82 to -20; p=0.0012) from baseline (table 3). Fitbit and charity participants showed similar step counts at months 6 and 12. However, MVPA bout min per week were significantly greater than baseline for the full sample among Fitbit participants (16 min [95% CI 2 to 30; p=0.0301]) and for the insufficiently active subset of charity participants (17 min [2 to 32; p=0.0231]). Cash participants showed MVPA bouts and steps similar to baseline.

We identified no evidence that provision of charity incentives produced higher MVPA bout min per week than provision of Fitbit only at the 12 month follow-up (table 3). Although the effects on steps in the active and insufficiently active participants in the Fitbit and charity groups at month 12 are similar to those at month 6 compared with control, largely because of the continued reduction in MVPA bout min per week in participants in the control group, the between-group comparison estimates of MVPA bout min per week are significantly greater in these groups compared with control, with overall differences exceeding 30 MVPA bout min per week in both groups. By contrast, although at month 12 the cash group still had a significantly higher number of average daily steps and a higher likelihood of meeting the

weekly step target than the control group, the increase in MVPA bout min per week was no longer significant. Moreover, the cash group did not have a higher number of average daily steps or a greater likelihood of reaching the weekly step target than the Fitbit or charity groups. In fact, we noted a potential undermining effect of cash incentives on MVPA bout min when comparing the cash group with the Fitbit group, driven by a net difference of 42 MVPA bout min per week (95% CI -79 to -5; p=0.0246) among participants who were already active at baseline.

None of the activity changes yielded significant improvements in health outcomes (such as BMI, systolic blood pressure, VO_{2max} , or EQ-5D-5L index) at either month 6 or month 12 (table 4). Results for all other secondary outcomes are shown in the appendix. Few secondary outcomes showed significant differences.

Discussion

Wearable activity trackers are becoming increasingly popular. However, our results show that they are unlikely to be a panacea for rising rates of chronic disease. Although the trackers seem to have been effective at stemming a reduction in physical activity seen in participants in the control group at 12 months, we identified no evidence of improved health outcomes. The reason for this result might be because the Fitbit group participants did not

	Difference from baseline, by study group				Between-group comparisons					
	Control (n=201)	Fitbit (n=203)	Charity (n=199)	Cash (n=197)	Fitbit vs control	Charity vs control	Charity vs Fitbit	Cash vs control	Cash vs Fitbit	Cash vs charity
Changes at month 6										
Weight (kg)	-0.3 (-1.0 to 0.4; p=0.3725)	0.2 (-0.6 to 0.9; p=0.7959)	0.1 (-0.6 to 0.9; p=0.8031)	0.4 (-0.3 to 1.1; p=0.3122)	0.49 (-0.54 to 1.53; p=0.4185)	0.45 (-0.59 to 1.49; p=0.4224)	-0.05 (-1.11 to 1.01; p=0.9950)	0.71 (-0.31 to 1.73; p=0.1805)	0.22 (-0.81 to 1.25; p=0.6063)	0.26 (-0.80 to 1.32; p=0.6003)
Systolic blood pressure (mm Hg)	2.5 (-0.8 to 5.9; p=0.1620)	-1.3 (-4.7 to 2.1; p=0.4838)	-0.5 (-4.0 to 3.0; p=0.7851)	-0.3 (-3.7 to 3.0; p=0.7488)	-3.80 (-8.09 to 0.49; p=0.0918)	-2.97 (-7.41 to 1.46; p=0.1833)	0.83 (-3.59 to 5.25; p=0.7423)	-2.85 (-7.05 to 1.35; p=0.1684)	0.95 (-3.28 to 5.18; p=0.7587)	0.12 (-4.23 to 4.47; p=0.9751)
NET-F VO _{2max} (mL/kg per min)	0.4 (-0.1 to 0.8; p=0.0809)	0.7 (0.3 to 1.1; p=0.0002)	0.5 (0.0 to 0.9; p=0.0107)	0.9 (0.5 to 1.3; p<0.0001)	0.35 (-0.21 to 0.92; p=0.1452)	0.10 (-0.50 to 0.69; p=0.5502)	-0.26 (-0.82 to 0.31; p=0.4138)	0.51 (-0.06 to 1.08; p=0.0778)	0.16 (-0.40 to 0.71; p=0.7523)	0.41 (-0.17 to 1.00; p=0.2577)
Change in EQ-5D-5L index	-2.1 (-4.0 to -0.1; p=0.0515)	-1.8 (-3.8 to 0.2; p=0.0863)	-2.1 (-4.1 to -0.1; p=0.0601)	-0.5 (-2.5 to 1.4; p=0.5731)	0.26 (-2.45 to 2.96; p=0.8517)	-0.03 (-2.72 to 2.65; p=0.9385)	-0.29 (-3.05 to 2.48; p=0.9144)	1.52 (-1.14 to 4.17; p=0.2861)	1.26 (-1.42 to 3.95; p=0.3844)	1.55 (-1.16 to 4.26; p=0.3273)
Changes at month 12										
Weight (kg)	-1.3 (-2.8 to 0.2; p=0.0957)	-0.4 (-2.1 to 1.2; p=0.5882)	-0.8 (-2.4 to 0.7; p=0.2777)	-0.6 (-2.1 to 1.0; p=0.4575)	0.9 (-1.2 to 3.0; p=0.4188)	0.5 (-1.5 to 2.4; p=0.6429)	-0.4 (-2.5 to 1.7; p=0.7041)	0.7 (-1.2 to 2.7; p=0.4665)	-0.1 (-2.2 to 1.9; p=0.8908)	0.3 (-1.7 to 2.3; p=0.7997)
Systolic blood pressure (mm Hg)	-0.3 (-3.9 to 3.3; p=0.8798)	-0.2 (-4.1 to 3.7; p=0.9319)	2.2 (-1.7 to 6.1; p=0.2619)	-0.9 (-4.6 to 2.8; p=0.6372)	0.1 (-4.9 to 5.1; p=0.9666)	2.5 (-2.3 to 7.3; p=0.3118)	2.4 (-2.7 to 7.4; p=0.3586)	-0.6 (-5.5 to 4.3; p=0.8057)	-0.7 (-5.8 to 4.4; p=0.7823)	-3.1 (-8.2 to 2.0; p=0.2369)
NET-F VO _{2max} (mL/kg per min)	1.4 (0.5 to 2.4; p=0.0031)	1.6 (0.6 to 2.6; p=0.0017)	1.3 (0.4 to 2.2; p=0.0061)	1.7 (0.7 to 2.6; p=0.0008)	0.2 (-0.8 to 1.2; p=0.7453)	-0.1 (-1.1 to 0.9; p=0.7767)	-0.3 (-1.3 to 0.7; p=0.5508)	0.2 (-0.8 to 1.2; p=0.6578)	0.1 (-1.0 to 1.1; p=0.9090)	0.4 (-0.6 to 1.4; p=0.4699)
Change in EQ-5D-5L index	-1.7 (-3.7 to 0.3; p=0.0898)	-1.3 (-3.3 to 0.6; p=0.1848)	-1.8 (-3.7 to 0.2; p=0.0722)	-0.1 (-2.0 to 1.8; p=0.9405)	0.4 (-2.2 to 3.0; p=0.7614)	-0.1 (-2.7 to 2.5; p=0.9626)	-0.5 (-3.2 to 2.2; p=0.7362)	1.6 (-1.0 to 4.3; p=0.2237)	1.2 (-1.5 to 4.0; p=0.3696)	1.7 (-0.9 to 4.3; p=0.2022)

Data are change in health outcomes (95% CI; p value). Comparisons across study groups of secondary outcomes at months 6 and 12 were adjusted for baseline outcome value, age, sex, and ethnic origin. NET-F=non-exercise testing for cardiorespiratory fitness. VO_{2max}=maximum oxygen consumption.

Table 4: Changes in health outcomes at months 6 and 12

show an increase in steps, and the increase in MVPA bout min relative to baseline was only 16 min per week on average, which is probably not enough to generate noticeable improvements in any health outcomes.

In this study, even among a selected group of employees who volunteered to participate, 40% abandoned the Fitbit within 6 months, and 90% did so by month 12. Because there seemed to be some beneficial effects in participants in the Fitbit group at month 12 despite the lack of wear, it is possible that participants wore the unit for a brief period of time, learned about their activity patterns, and then stopped wearing the device. This knowledge might have allowed these participants to stem the reduction in physical activity that was seen in the control group, despite the lack of continued wear.

Although participants in the charity group earned a mean of S\$320.00 for charity, most likely by increasing the amount of time they wore the device, this strategy does not seem to be a good use of resources with respect to the employee's physical activity. Results, in terms of physical activity and health outcomes with the charity incentive, were no better than just providing the participant with a Fitbit device.

The results for the cash group are not so straightforward. Cash incentives clearly affected participant behaviour.

Unlike the other groups, during the incentive period, cash motivated both insufficiently active and active participants to increase physical activity levels. These results are consistent with findings from three other randomised controlled trials²⁸⁻³⁰ that tested the use of incentives to influence physical activity, albeit using incentive schemes very different from that used in the present study. As with our study, two of the three studies showed significant increases in activity, although none was longer than 16 weeks in duration.

Despite the step effect, the net effect of the cash incentive on MVPA bouts was small, averaging just more than an additional 4 min per day. Consequently, the fact that the step increases did not translate into noticeable increases in health outcomes is not surprising. Despite this small effect on activity, the mean earnings of participants in the cash group was S\$620.00. This cash incentive is a substantial expense and one that most third-party payers would be unlikely to undertake without evidence of improvements in health outcomes, especially considering that once the incentives were removed activity levels trended back to baseline values.

The absence of improvements in health outcomes in the cash group might be related to how the cash incentive was implemented. We chose to focus participants on

meeting weekly step targets and used somewhat arbitrary targets and reward amounts. The strategy used in this study is one among a myriad of potential strategies. Although it is possible that our study was too short to show improvements in the chosen health indicators or that the population was too healthy at baseline for the intervention to produce a significant difference, the absence of an improvement in health outcomes might be due to the small increase in MVPA bouts min in participants assigned to the cash group. On average, cash incentive participants increased their weekly steps by 7336, or a distance of roughly 5.9 km, between baseline and month 6. However, this increase in steps translated into only a small increase of 4 min per day in MVPA bouts. The rewards might have been better spent had they targeted MVPA min or MVPA steps (sometimes called active or aerobic steps) directly, which many activity trackers also report. However, in view of the absence of increases in activity or improvements in health outcomes at month 12, any incentive strategy would probably need to be sustained over a longer period of time for health improvements to be seen.

This study has several limitations. Because the sample comprised full-time employees from selected worksites who volunteered to participate, our participants were probably healthier and more motivated to be physically active than the average full-time worker. The reduction in physical activity reported in control participants might be because their baseline physical activity was artificially high as a result of joining the study or wearing the accelerometer for the first time. For these reasons, our results might not have external validity. Another issue is that of potential contamination because participants within worksites but in different study groups might have interacted with each other. Although we would have preferred to randomise at the worksite level, we were unable to convince employers to participate in a study where some worksites might have been randomly assigned to the control group. Consequently, our results might be conservative because contamination would bias the results towards the null hypothesis.

In summary, our results provide some evidence of the ability of the tracker to improve step activity. However, we find no evidence of improvements in health outcomes, either with or without incentives. Since cash incentives were shown to increase weekly steps, other incentive strategies that provide a stronger link between the target of the incentives and health outcomes, such as incentivising MVPA bout min directly, might yield greater effectiveness. However, results suggest any cash incentive strategy would probably need to be in place long term to avoid any potential undermining effect resulting from their removal.

Contributors

EAF was the principal investigator. EAF, RAS, and MB were responsible for the study design. EAF and AS were responsible for the study protocol and data gathering. BAH and EEKN were responsible for data analysis.

KRE helped develop protocols for accelerometer data collection and analyses. All authors contributed to interpretation of the results. EAF and AS prepared the first draft of the reports, and all authors critically read and edited several drafts before submission. All authors read and approved the submitted version.

Declaration of interests

We declare no competing interests.

Acknowledgments

This study was supported by a Health Services Research Competitive Research Grant (HSRG/0022/2012) from the Ministry of Health, Singapore.

References

- 1 WHO. Global action plan for the prevention and control of noncommunicable diseases 2013–2020. Geneva: World Health Organization, 2013.
- 2 Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet* 2012; **380**: 219–29.
- 3 Finkelstein EA, daCosta DiBonaventura M, Burgess SM, Hale BC. The costs of obesity in the workplace. *J Occup Environ Med* 2010; **52**: 971–76.
- 4 Ledger D, McCaffrey D. Inside wearables: how the science of human behaviour change offers the secret to long-term engagement. Endeavour Partners, 2014. <http://endeavourpartners.net/assets/Endeavour-Partners-Wearables-White-Paper-2014L.pdf> (accessed Sept 6, 2016).
- 5 Zulman DM, Damschroder LJ, Smith RG, et al. Implementation and evaluation of an incentivized Internet-mediated walking program for obese adults. *Transl Behav Med* 2013; **3**: 357–69.
- 6 Chevron US. Benefits 2015. http://hr2.chevron.com/usbenefits/cardio/healthevent_pedom.asp (accessed May 16, 2016).
- 7 Bravata DM, Smith-Spangler C, Sundaram V, et al. Using pedometers to increase physical activity and improve health: a systematic review. *JAMA* 2007; **298**: 2296–304.
- 8 Freak-Poli R, Cumpston M, Peeters A, Clemes SA. Workplace pedometer interventions for increasing physical activity. *Cochrane Database Syst Rev* 2013; **4**: CD009209.
- 9 Finkelstein EA, Sahasranaman A, John G, et al. Design and baseline characteristics of participants in the trial of economic incentives to promote physical activity (TRIPPA): a randomized controlled trial of a six month pedometer program with financial incentives. *Contemp Clin Trials* 2015; **41**: 238–47.
- 10 Ajzen I, Fishbein M. Understanding attitudes and predicting social behaviour. London: Pearson, 1980.
- 11 Ariely D, Jones S. Predictably irrational. New York: HarperCollins, 2008.
- 12 Singapore Ministry of Health. National health survey 2010. Singapore: Epidemiology & Disease Control Division Ministry of Health, 2011.
- 13 Thomas S, Reading J, Shephard RJ. Revision of the physical activity readiness questionnaire (PAR-Q). *Can J Sport Sci* 1992; **17**: 338–45.
- 14 Singapore Health Promotion Board. Active for life [physical activity: some is better than none]. 2011. http://www.hpb.gov.sg/HOPPortal/content/conn/HOPUCM/path/Contribution%20Folders/uploadedFiles/HPB_Online/Health_Topics/Physical_Activity/ActiveForLife%20booklet_20110415_final.pdf (accessed Sept 16, 2016).
- 15 Singapore Health Promotion Board. Brisk walk your way to better health. 2011. http://www.hpb.gov.sg/HOPPortal/content/conn/HOPUCM/path/Contribution%20Folders/uploadedFiles/HPB_Online/Educational_Materials/HPB%20Briskwalking%20Booklet_Eng.pdf (accessed Sept 16, 2016).
- 16 Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. *Int J Behav Nutr Phys Act* 2015; **12**: 159.
- 17 Domelen DV. Functions for processing minute-to-minute accelerometer data version 2.2.2. <http://cran.r-project.org/web/packages/accelerometry/index.html> (accessed Oct 3, 2014).
- 18 Sasaki JE, John D, Freedson PS. Validation and comparison of ActiGraph activity monitors. *J Sci Med Sport* 2011; **14**: 411–16.
- 19 Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc* 2008; **40**: 181–88.

- 20 Sloan RA, Haaland BA, Leung C, Padmanabhan U, Koh HC, Zee A. Cross-validation of a non-exercise measure for cardiorespiratory fitness in Singaporean adults. *Singapore Med J* 2013; **54**: 576–80.
- 21 Group TE. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy* 1990; **16**: 199–208.
- 22 Tongsir S, Cairns J. Estimating population-based values for EQ-5D health states in Thailand. *Value Health* 2011; **14**: 1142–45.
- 23 Strath SJ, Holleman RG, Ronis DL, Swartz AM, Richardson CR. Objective physical activity accumulation in bouts and nonbouts and relation to markers of obesity in US adults. *Prev Chronic Dis* 2008; **5**: A131.
- 24 Cohen J. *Statistical power analysis for the behavioral sciences*. Oxford, UK: Routledge Academic, 2013.
- 25 Gulliford MC, Ukoumunne OC, Chinn S. Components of variance and intraclass correlations for the design of community-based surveys and intervention studies: data from the Health Survey for England 1994. *Am J Epidemiol* 1999; **149**: 876–83.
- 26 Li K-H, Raghunathan TE, Rubin DB. Large-sample significance levels from multiply imputed data using moment-based statistics and an F reference distribution. *J Am Stat Assoc* 1991; **86**: 1065–73.
- 27 Rubin DB. *Multiple imputation for nonresponse in surveys*. Hoboken, NJ, USA: John Wiley & Sons, 2004.
- 28 Finkelstein EA, Brown DS, Brown DR, Buchner DM. A randomized study of financial incentives to increase physical activity among sedentary older adults. *Prev Med* 2008; **47**: 182–87.
- 29 Petry NM, Andrade LF, Barry D, Byrne S. A randomized study of reinforcing ambulatory exercise in older adults. *Psychol Aging* 2013; **28**: 1164–73.
- 30 Kullgren JT, Harkins KA, Bellamy SL, et al. A mixed-methods randomized controlled trial of financial incentives and peer networks to promote walking among older adults. *Health Educ Behav* 2014; **41** (suppl 1): 43S–50S.