Arthritis Care & Research
Vol. 70, No. 5, May 2018, pp 732–740
DOI 10.1002/acr.23324

ORIGINAL ARTICLE

Financial Incentives and Health Coaching to Improve Physical Activity Following Total Knee Replacement: A Randomized Controlled Trial

ELENA LOSINA,[1] JAMIE E. COLLINS,[2] BHUSHAN R. DESHPANDE,[2] SAVANNAH R. SMITH,[3] GRIFFIN L. MICHL,[3] ILANA M. USISKIN,[3] KRISTINA M. KLARA,[3] AMELIA R. WINTER,[3] HEIDI Y. YANG,[3] FAITH SELZER,[2] AND JEFFREY N. KATZ[4]

Objective. **Most persons who undergo total knee replacement (TKR) do not increase their physical activity following surgery. We assessed whether financial incentives and health coaching would improve physical activity in persons undergoing TKR.**
Methods. **We designed a factorial randomized controlled trial among persons undergoing TKR for osteoarthritis. Subjects underwent normal perioperative procedures, including postoperative physical therapy, and were assigned to 1 of 4 arms: attention control, telephonic health coaching (THC), financial incentives (FI), or THC + FI. We objectively measured step counts and minutes of physical activity using a commercial accelerometer (Fitbit Zip) and compared the changes from pre-TKR to 6 months post-TKR across the 4 study arms.**
Results. **Of the 202 randomized subjects, 150 (74%) provided both pre-TKR and 6 months post-TKR accelerometer data. Among completers, the mean $\pm$ SE daily step count at 6 months ranged from 5,619 $\pm$ 381 in the THC arm to 7,152 $\pm$ 407 in the THC + FI arm (adjusting for baseline values). Daily step count 6 months post-TKR increased by 680 (95% confidence interval [95% CI] −94, 1,454) in the control arm, 274 (95% CI −473, 1,021) in the THC arm, 826 (95% CI 89, 1,563) in the FI arm, and 1,808 (95% CI 1,010, 2,606) in the THC + FI arm. Weekly physical activity increased by mean $\pm$ SE 14 $\pm$ 10, 14 $\pm$ 10, 16 $\pm$ 10, and 39 $\pm$ 11 minutes in the control, THC, FI, and THC + FI arms, respectively.**
Conclusion. **A dual THC + FI intervention led to substantial improvements in step count and physical activity following TKR.**

INTRODUCTION

Regular engagement in physical activity (PA) is critically important for the prevention of chronic diseases and improvement of overall health (1,2). Even relatively modest increases in PA have been shown to markedly reduce the risk of mortality (3). PA is particularly important for persons with knee osteoarthritis (OA) given its ability to address pain and functional limitations (4). However, adherence to PA guidelines by persons with knee OA is poor (5).

Half of all individuals diagnosed with knee OA eventually choose to undergo total knee replacement (TKR) (6). While TKR results in dramatic improvements in pain and functional capabilities in more than 80% of patients (7), there is

A video abstract of this article can be found at https://bcove.video/2vRtNUS and https://vimeo.com/265584038.
ClinicalTrials.gov identifier: NCT01970631.
Supported by the NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases (grants R21-AR-063913 and K24-AR-057827).
[1]Elena Losina, PhD: Orthopedic and Arthritis Center for Outcomes Research and the Policy and Innovation Evaluation in Orthopedic Treatments Center, Brigham and Women's Hospital, Harvard Medical School, and Boston University School of Public Health, Boston, Massachusetts; [2]Jamie E. Collins, PhD, Bhushan R. Deshpande, BS, Faith Selzer, PhD: Orthopedic and Arthritis Center for Outcomes Research and the Policy and Innovation Evaluation in Orthopedic Treatments Center, Brigham and Women's Hospital, and Harvard Medical School, Boston, Massachusetts; [3]Savannah R. Smith, BA (current address: George Washington University School of Medicine & Health Sciences, Washington, DC), Griffin L. Michl, BS (current address: Reddit, San Francisco, California), Ilana M. Usiskin, BS (current address: Yale School of Medicine, New Haven, Connecticut), Kristina M. Klara, BS (current address: Yale School of Medicine, New Haven, Connecticut), Amelia R. Winter, BA, Heidi Y. Yang, MS, MPH: Orthopedic and Arthritis Center for Outcomes Research and the Policy and Innovation Evaluation in Orthopedic Treatments Center, Brigham and Women's Hospital, Boston, Massachusetts; [4]Jeffrey N. Katz, MD, MSc: Orthopedic and Arthritis Center for Outcomes Research and the Policy and Innovation Evaluation in Orthopedic Treatments Center, Brigham and Women's Hospital, Harvard Medical School, and Harvard T. H. Chan School of Public Health, Boston, Massachusetts.
Drs. Losina and Katz are Deputy Editors for the *Journal of Bone and Joint Surgery*. Dr. Katz is President-Elect of the Osteoarthritis Research Society International and a medical editor for Healthwise. Dr. Losina is a statistical consultant for TissueGene.
Address correspondence to Elena Losina, PhD, Department of Orthopedic Surgery, Brigham and Women's Hospital, 60 Fenwood Road, Suite 5016-U, Boston, MA 02115. E-mail: elosina@bwh.harvard.edu.
Submitted for publication January 30, 2017; accepted in revised form July 11, 2017.

Significance & Innovations

- Physical activity levels are normally unchanged following total knee replacement.
- Commercially available accelerometers allow for objective assessment of physical activity.
- Financial incentives integrated with remote health coaching for improving physical activity is a previously untested intervention among patients who are receiving total knee replacement.
- We randomly assigned recipients of total knee replacement to attention control, telephonic health coaching, financial incentives, or both interventions. Compared to attention control, the combined intervention led to substantial improvements in both daily steps and moderate-to-vigorous physical activity.

typically minimal change in PA levels following surgery (8,9). The relative lack of change could be explained by a variety of reasons. Many patients are sedentary preoperatively, which is a strong predictor of being sedentary postoperatively, suggesting that individuals may have to substantially alter their lifestyles to increase physical activity post-TKR (10). Especially in the period immediately following surgery, patients may experience pain with exercise. Patients with overly optimistic preoperative expectations for their functional and pain trajectories may be reluctant to pursue PA further (11). However, the postoperative period, which is marked by intensive engagement with physical therapy and multidisciplinary clinical care, offers a unique opportunity to change patients' attitudes and regular behaviors regarding PA.

We investigated 2 interventions that held promise for increasing PA among TKR recipients: telephonic health coaching and financial incentives. Health coaching interventions have been proven efficacious in improving self-efficacy, medication adherence, and health-related outcomes in patients with chronic diseases (12–14). Financial incentives employ techniques from behavioral economics, addressing the fact that many individuals disproportionately prefer immediate gratification rather than delayed rewards (15–18). We hypothesized that health coaching could facilitate internal motivation to increase engagement in walking and other age-appropriate physical activities and that a rewards-based intervention could offer external motivation to incentivize TKR recipients to strive for the long-term benefits of physical activity. We designed and conducted a randomized controlled trial utilizing a 2x2 factorial design to evaluate the efficacy of financial incentives and health coaching to improve PA following TKR.

PATIENTS AND METHODS

Screening, enrollment, and randomization. The Study of Physical Activity Rewards Following Knee Surgery (SPARKS) was approved by the local institutional review board and was preregistered at ClinicalTrials.gov. The study screening and enrollment took place from November 2013 through January 2016. We screened patients with knee OA who were scheduled to undergo a primary, unilateral TKR by 1 of 5 orthopedic surgeons at Brigham and Women's Hospital, a tertiary medical center in Boston, Massachusetts. Participants were ineligible if they were younger than 40 years, did not speak English, resided in a nursing home, were scheduled to undergo a contralateral TKR or other surgery requiring hospitalization within 6 months, were previously diagnosed with an inflammatory arthritis or osteonecrosis affecting the knee, had a comorbidity that might prevent the safe performance of moderate ambulatory physical activity, including epilepsy, Parkinson's disease, neuropathy, or dementia, or who required a wheelchair or walker to ambulate preoperatively (19). We also excluded persons who did not have regular access to the internet, which was needed for the transfer of accelerometer data to the central server.

On the day of their preoperative appointment, patients provided written informed consent, underwent a brief musculoskeletal examination, and completed a 27-item questionnaire to assess their levels of "delay discounting" (20). Delay discounting is a measure of impulsivity, which evaluates how significantly individuals value long-term health benefits, such as those associated with physical activity. Participants completed a baseline questionnaire, which included demographic information, social and employment history, and resource utilization, as well as validated measures such as the Knee Injury Outcomes and Osteoarthritis Score (KOOS) (21), the 5-dimension EuroQol, a general health visual analog scale (22), the Risk Taking Index (23), the Work Productivity and Activity Impairment questionnaire (24,25), the Yale Physical Activity Survey (26,27), self-reported knee range of motion (28), and several components of the 36-item Short Form Health Survey (SF-36) (including the SF-12, the Mental Health Inventory, and the Vitality Score) (29). Subjects were given a commercial accelerometer (Fitbit Zip) and were required to wear it for at least 5 days within a 7-day period prior to surgery. The accelerometer measures the wearer's steps on a minute by minute basis. The Fitbit Zip was selected because it was a low-cost wearable device regarded as convenient and comfortable for weeklong wear by research subjects and was previously validated among younger adults for measuring step count and physical activity (30,31). Participants did not have access to the accelerometer's online account until the conclusion of the study to ensure that the accelerometer was used as a measuring device and as a motivating tool to view long-term trends; participants could, however, view steps taken each day they wore the Fitbit.

Following confirmation of the surgery's completion, participants were randomized into 1 of 4 groups: attention control, telephonic health coaching (THC), financial incentives (FI), or THC + FI. Randomization was implemented in block sizes of 4 and was stratified by age (<65 versus ≥65 years), preoperative average daily steps (<3,000 versus ≥3,000), and delay discounting (<0.00256 versus ≥0.00256, with higher values indicating a greater preference for immediate rewards) (32). Subjects were discontinued if they underwent any surgery requiring overnight hospitalization or if they requested to withdraw. To limit factors unrelated to the evaluated interventions that could affect PA, we discontinued study participants if they underwent a surgery requiring overnight hospitalization, because such surgeries are frequently associated with subsequent limited activity during the recovery period.

Interventions. *All arms.* Subjects were informed of their assigned group during their first phone call following surgery, and the interventions began the week following surgery. Study participants were discharged to either a rehabilitation facility or to their home, where they received in-home and subsequently outpatient physical therapy according to normal postoperative practice. Participants in each arm received calls on a regular basis; calls were made weekly for weeks 2–5 following TKR and biweekly for weeks 7–24, for a total of 14 instances.

Attention control. Participants in the attention control and FI arms received attention control calls. These calls conveyed general health messages; conversations focused on general aspects of recovery and rehabilitation. Staff making calls avoided motivational interviewing techniques and discussion of physical activity. If participants could not be reached, a voicemail was left reminding them that they could contact study staff should they have any questions about the study. Only the date and time of the call was recorded for attention control calls.

THC. The THC intervention consisted of periodic calls made by research staff trained in motivational interviewing techniques. The health coaches reviewed materials written by Miller and Rollnick and received case-based training with detailed formal analysis by senior investigators with experience in motivational interviewing (33,34). Additionally, health coaches attended the local preoperative education class offered to TKR recipients and observed both surgeries and postoperative physical therapy sessions to understand the process from a patient's perspective. Health coaches took notes for each call, including content discussed and the length of each call and regularly debriefed coaching technique with the study clinical investigator (JNK).

The THC calls were intended to focus on the subjects' short- and long-term PA goals. Rather than imposing goals on study subjects, health coaches used open-ended questions to elicit the participants' own objectives, while expressing empathy, recognizing and helping to resolve any discrepancy between subjects' PA goals and current behavior, rolling with resistance, and supporting self-efficacy (33). Coaches called at the times identified as most convenient by the study participants and made up to 4 call attempts at each time point.

FI. The FI program was designed to be implemented in a timely and consistent manner and to provide an opportunity for subjects to receive both smaller and larger rewards (35). Participants randomized to the FI arms were assigned an escrow account initially containing $105. Participants logged on to the study website to complete daily PA logs for postoperative weeks 2–8 and weekly physical activity logs for weeks 9–23 (except week 12). Participants received $5 from their escrow account for completing at least 5 of 7 daily logs or the 1 weekly log. Throughout the study, subjects were able to log on to the study website to view the amount of money they "lost" due to missed reporting. These rewards were designed to develop the discipline of logging the exercises performed during the post-TKR rehabilitation period.

During weeks 14–23, subjects were eligible to receive a $15 bonus payment for every week that they increased their self-reported minutes of PA by at least 10% from the preceding week, as calculated from part 1 of the Yale Physical Activity Scale (26). At 6 months, subjects in the FI arms were eligible to receive an additional $50 payment if they either increased their daily step count by at least 50% compared to 3 months previously or met a PA guideline of ≥150 minutes of moderate-to-vigorous PA (MVPA) per week. In total, subjects were eligible to earn a maximum of $305 from the FI component over the study duration. Checks were mailed to participants every other week.

Followup. All subjects were asked to complete study questionnaires at 3 and 6 months post-TKR. They also wore their accelerometer for 1 week at each of these time points. For each time point, participants were paid $25 for completing the questionnaire and $5 for each day that they wore their accelerometer. Study activities and followup for each randomization group are outlined in Supplementary Figure 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23324/abstract.

Outcomes. The primary outcome was predefined as the mean number of steps/day at 6 months post-TKR. Secondary outcome measures included the change in the mean number of steps/day between baseline and 6 months and change in weekly minutes of MVPA from baseline to 6 months.

We downloaded participants' minute-by-minute step count data through Fitbit's application program interface. Because the Fitbit Zip measures and outputs step counts rather than intensity counts, we relaxed the typical accelerometer-based thresholds of 10 hours of wear time and 60 minutes of nonwear (36), defining valid days of data as days with at least 8 hours of wear time, and nonwear time was defined as periods ≥90 minutes during which no steps were recorded. We calculated an average daily step count over the days of valid Fitbit wear. In the primary analysis, we included participants who provided accelerometer data for at least 4 valid days in a 7-day period at both baseline and 6 months.

To derive the MVPA duration, we calculated bouts during which the study participant took ≥100 steps/minute during a Fitbit valid day. We required that bouts be at least 10 minutes long, in accordance with PA guidelines (37), and within a single bout we allowed up to 2 grace minutes, during which the participant's step count could fall below the 100 steps/minute threshold. The 100 steps/minute threshold has previously been shown to correspond to an energy expenditure of 3 metabolic equivalents of task (METs) and has been used in analyses of older adults (38,39).

Statistical analysis. Prior data suggest that the efficacy of each intervention is consistent with an effect size of 0.3 SDs (35,40,41). We hypothesized that the 2 interventions together would have an efficacy consistent with an effect size of 0.6 SD. Using a *t*-test and assuming a 2-sided alpha (Type I error rate) of 5%, statistical power of 80%, and 15% loss to followup, we estimated a sample size of 50 subjects per arm (200 total) (42).

We compared the baseline demographic and clinical characteristics of participants across each of the 4 arms to ensure that the measured covariates were balanced across arms. We used general linear modeling to evaluate the difference in primary and secondary outcomes between the 4 arms, after adjustment for imbalanced baseline covariates. In addition, we examined the difference between the control arm and each of the intervention arms, assessing improvements in step count and MVPA for clinical relevance using the effect size. Effect sizes are defined as a ratio of the difference between 2 groups relative to the SD. The effect size emphasizes the size of the effect and is not affected by the sample size; an effect size of 0.35–0.5 SD is generally considered clinically relevant (43).

In the primary analysis, we included participants who provided accelerometer data for at least 4 valid days in a 7-day period both at baseline and at 6 months. In the sensitivity analysis, we augmented the missing data at 6 months in the following manner: for those participants providing between 1 and 3 valid days, we calculated daily step count and minutes of MVPA from the available data, and for those with 0 valid days, we used multiple imputation techniques to augment missing 6 months PA and step count data. We repeated the sensitivity analysis, excluding participants who were discontinued from the study prior to 6 months. Participants not providing any valid days of data at baseline were excluded from all analyses. Data analyses were conducted using SAS, version 9.4.

RESULTS

Study sample. We identified 1,245 individuals scheduled to undergo TKR, of whom 632 were eligible and contacted for recruitment. Of those, 258 individuals agreed to participate in the study. Prior to the planned date of surgery, 47 persons who intended to participate did not complete the online baseline questionnaire or wore the accelerometer for fewer than 5 days. Nine persons chose to not undergo surgery, leaving 202 persons who were randomized and comprised our study sample (Figure 1).

Randomized participants had a mean ± SD age of 65 ± 8 years; 57% were female; 68% had at least a bachelor's degree. The cohort's mean ± SD body mass index was 31 ± 6 kg/m^2. The mean ± SD KOOS pain score prior to TKR was 53 ± 18, and the mean ± SD KOOS function score was 59 ± 18 (range 0–100, where 100 is worst). Prior to TKR, participants walked

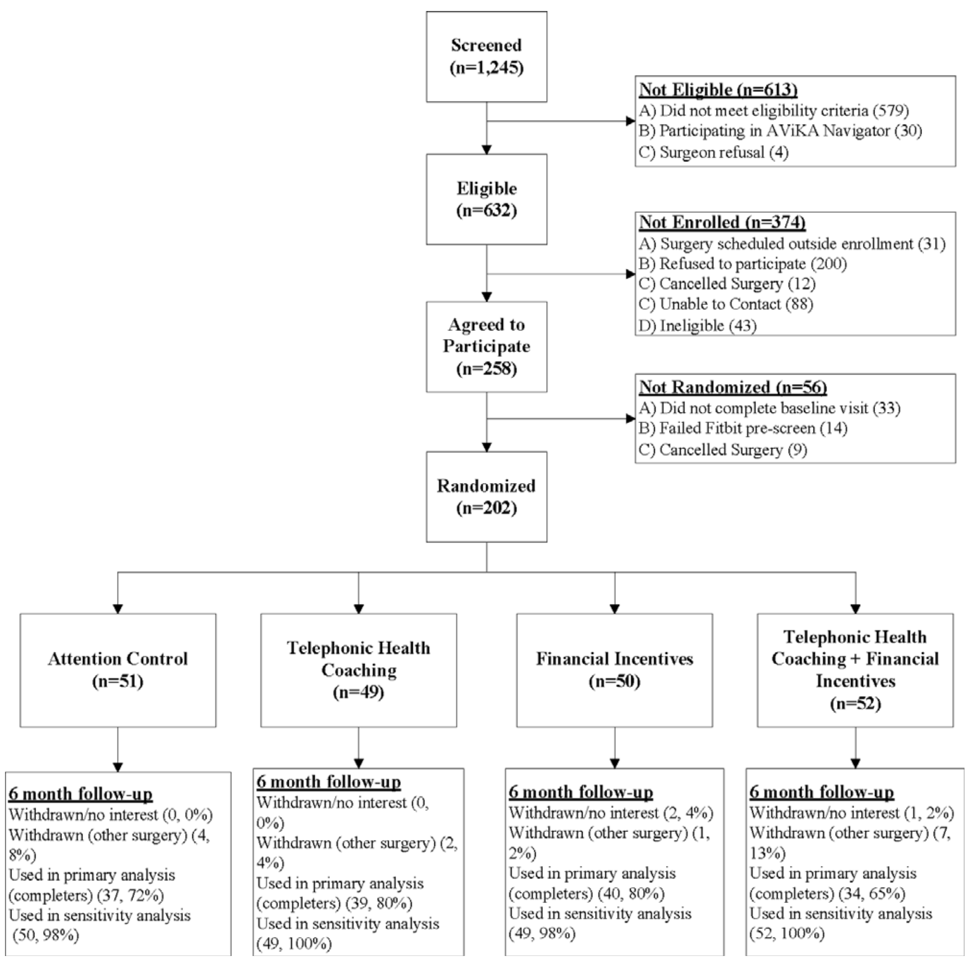

Figure 1. Study enrollment and followup (CONSORT [Consolidated Standards of Reporting Trials] diagram). A total of 1,245 persons were assessed for eligibility, of whom 632 were deemed eligible for enrollment. Of the 258 eligible who agreed to participate, 202 were ultimately randomized to 1 of the 4 study arms.

a mean ± SD 5,032 ± 2,771 steps/day. Baseline characteristics by study arm are shown in Table 1; with the exception of step count, all baseline characteristics are balanced across the arms (Table 2 and Table 3). Eligible patients who refused to participate in the study were somewhat older (69 years old, on average) and were more likely to be female (63%), but these differences were not large.

Adherence to intervention protocol. Study staff made 1,325 (97% of possible 1,372) attention control calls to active study participants randomized to receive such calls. Forty-six percent of these calls resulted in reaching the subject; the remainder led to voice messages. In the THC arms, the mean ± SD rate of completed calls was 72% ± 22% across 14 time points over 6 months. Eighty-six percent of subjects received at least 7 of the required 14 calls, while half received ≥10 calls. The average call time varied from 18 minutes for the first call (1 week post-TKR) to 11–15 minutes for subsequent weeks.

Study participants in the FI arms received a total of 73% of possible payments from their escrow accounts for completing daily activity logs and 76% for weekly activity logs. They earned 38% of the potential weekly $15 bonuses for increasing their PA by at least 10% from the preceding

Table 1. Baseline characteristics of the study sample by arm*

Characteristic	AC	THC	FI	FI + THC
Sample size, no.	51	49	50	52
Age, mean ± SD years	65.8 ± 6.9	65.0 ± 6.9	65.0 ± 8.3	65.7 ± 8.1
Sex				
Male	24 (47.1)	19 (38.8)	17 (34.0)	27 (51.9)
Female	27 (52.9)	30 (61.2)	33 (66.0)	25 (48.1)
Education				
High school or less	6 (11.8)	7 (14.3)	3 (6.0)	3 (5.8)
Some college	9 (17.6)	12 (24.5)	11 (22.0)	13 (25.0)
≥ bachelor's degree	36 (70.6)	30 (61.2)	36 (72.0)	36 (69.2)
Race†				
White	46 (90.2)	44 (89.8)	44 (88.0)	46 (90.2)
Black	2 (3.9)	2 (4.1)	4 (8.0)	2 (3.9)
Asian	0 (0)	0 (0)	0 (0)	2 (3.9)
Native American	1 (2.0)	0 (0)	0 (0)	0 (0)
Other	2 (3.9)	3 (6.1)	2 (4.0)	1 (2.0)
Annual household income				
<$29,900	2 (4.1)	2 (4.3)	5 (10.4)	4 (7.7)
$30,000–59,999	10 (20.4)	8 (17.0)	8 (16.7)	8 (15.4)
$60,000–99,999	15 (30.6)	17 (36.2)	12 (25.0)	15 (28.8)
>$100,000	22 (44.9)	20 (42.5)	23 (47.9)	25 (48.1)
Employment				
Full time	17 (34.7)	21 (43.7)	21 (42.0)	16 (32.0)
Part time	11 (22.4)	7 (14.6)	10 (20.0)	5 (10.0)
Not working	21 (42.9)	20 (41.7)	19 (38.0)	29 (58.0)
Body mass index, kg/m^2				
<30.0	22 (43.1)	19 (38.8)	22 (44.0)	29 (55.8)
30.0–34.9	17 (33.3)	16 (32.7)	11 (22.0)	13 (25.0)
≥35.0	12 (23.5)	14 (28.6)	17 (34.0)	10 (19.2)
Geriatric comorbidities‡				
No geriatric problems	25 (50.0)	23 (47.9)	17 (34.0)	21 (41.2)
1 geriatric problem	11 (22.0)	16 (33.3)	15 (30.0)	10 (19.6)
≥2 geriatric problems	14 (28.0)	9 (18.7)	18 (36.0)	20 (39.2)
HRQoL, mean ± SD§	0.7 ± 0.1	0.7 ± 0.1	0.7 ± 0.2	0.7 ± 0.1
Pain and physical function, mean ± SD				
KOOS pain	51.2 ± 17.4	56.6 ± 16.7	52.1 ± 20.6	53.2 ± 17.1
KOOS function	55.8 ± 18.7	61.7 ± 15.5	58.2 ± 17.7	59.9 ± 18.4
KOOS sports and recreation	21.9 ± 21.6	25.2 ± 21.7	26.3 ± 25.4	29.3 ± 23.1
SF-36, mean ± SD				
MHI-5	81.1 ± 14.3	80.0 ± 13.3	77.2 ± 16.9	81.8 ± 13.8
Range of motion (passive)				
Extension (>5° from straight)	34 (66.7)	31 (63.3)	23 (46.0)	29 (56.9)
Flexion <120°	21 (41.2)	21 (42.9)	23 (46.0)	22 (42.3)

* Values are the number (%) unless otherwise indicated. AC = attention control; THC = telehealth coaching; FI = financial incentives; HRQoL = health-related quality of life; KOOS = Knee Injury and Osteoarthritis Outcome Score; SF-36 = Short Form 36 health survey; MHI-5 = Mental Health Inventory.
† May sum to greater than 100%, as participants were allowed to select multiple options.
‡ Included problems with vision and hearing, falling, and incontinence.
§ Derived from the EuroQol 5-domain instrument.

Table 2. Followup rates overall and by arm*

Characteristic	AC	THC	FI	FI + THC
Baseline				
Accelerometer worn ≥4 valid days	50 (98.0)	46 (93.9)	49 (93.9)	47 (90.4)
6 months				
Accelerometer worn ≥4 valid days	38 (74.5)	39 (79.6)	41 (82.0)	35 (67.3)
Accelerometer worn 1–3 valid days	6 (11.8)	6 (12.2)	3 (6.0)	2 (3.8)

* Values are the number (%) unless otherwise indicated. AC = attention control; THC = telehealth coaching; FI = financial incentives.

week. At 6 months, 21% earned the $50 bonus for increasing their step count by at least 50% from month 3 or by meeting PA guidelines.

Outcomes. Of the 200 (of 202) randomized subjects who provided any baseline accelerometer data, 192 wore the accelerometer for at least 4 valid days at baseline. By 6 months, 17 subjects had withdrawn and 33 provided suboptimal accelerometer data, leaving 150 randomized subjects (74%) who wore the accelerometer for at least 4 valid days at both time points and comprised the analytic cohort for the complete case analysis. The mean ± SE daily step count at baseline (pre-TKR) varied across the study arms, with the highest being 6,158 ± 427 in the control arm; the rest of the arms had similar step counts, ranging from 4,974 to 5,229. MVPA also varied across the study arms, with the highest being 22 ± 5 in the control arm and the lowest being 9 ± 5 minutes in the THC arm (Table 3).

After adjustment for the baseline step count, the mean ± SE daily step count at 6 months ranged from 5,619 ± 381 in the THC arm to 7,152 ± 407 in the THC + FI arm; the control arm (mean 6,025) and FI arm (mean 6,170) achieved similar step counts ($P = 0.0502$ overall). The mean increase in daily step count from baseline to 6 months was 680 (95% confidence interval [95% CI] −94, 1,454), 274 (95% CI −473, 1,021), 826 (95% CI 89, 1,563), and 1,808 (95% CI 1,010, 2,606) in the control, THC, FI, and THC + FI arms, respectively (Table 3) (Figure 2). The difference in the change in average daily steps over the 6-month period between the THC + FI arm and the control arm was 1,128 (95% CI 14, 2,241).

The mean change in weekly MVPA between baseline and 6 months for participants in the control arm was 14 minutes (95% CI −6, 34), for participants in the THC arm was 14 minutes (95% CI −6, 33), for participants in the FI arm was 16 minutes (95% CI −3, 35), and for participants in THC + FI arm was 39 minutes (95% CI 18, 60). The THC + FI arm had a 25-minute (95% CI −4, 54) greater change in MVPA compared to the control arm ($P = 0.0967$) (Table 3).

Sensitivity analysis. In the sensitivity analysis, we computed the average step count for the 17 subjects (8%) providing only 1–3 days of accelerometer data at either of the 2 time points and used multiple imputation to augment missing data for the 32 subjects who provided baseline but not 6-month accelerometer data. Two subjects were excluded

Table 3. Primary and secondary outcomes of the SPARKS trial*

Characteristic	AC	THC	FI	FI + THC	P
Sample size, no.	37	39	40	34	
Primary outcome: average daily steps					
Baseline	6,158 (5,320, 6,995)	4,974 (4,158, 5,790)	5,052 (4,246, 5,858)	5,229 (4,355, 6,103)	0.1789
6 months (crude)	6,712 (5,670, 7,755)	5,305 (4,290, 6,321)	5,923 (4,920, 6,926)	7,054 (5,967, 8,142)	0.0924
6 months (adjusted for baseline)	6,025 (5,251, 6,799)	5,619 (4,871, 6,366)	6,170 (5,433, 6,907)	7,152 (6,354, 7,951)	0.0502
Change, baseline to 6 months†	680 (−94, 1,454)	274 (−473, 1,021)	826 (89, 1,563)	1,808 (1,010, 2,606)	0.0502
Secondary outcomes: weekly PA					
Baseline	22 (12, 33)	9 (−1, 19)	14 (4, 24)	10 (−1, 20)	0.2526
6 months (crude)	35 (13, 57)	23 (2, 45)	30 (9, 51)	49 (26, 72)	0.4250
6 months (adjusted for baseline)	28 (8, 48)	28 (8, 47)	30 (11, 49)	52 (32, 73)	0.2667
Change, baseline to 6 months†	14 (−6, 34)	14 (−6, 33)	16 (−3, 35)	39 (18, 60)	0.2667

* Values are the mean (95% confidence interval) unless otherwise indicated. SPARKS = Study of Physical Activity Rewards Following Knee Surgery; AC = attention control; THC = telehealth coaching; FI = financial incentives; PA = physical activity.
† Adjusted for baseline.

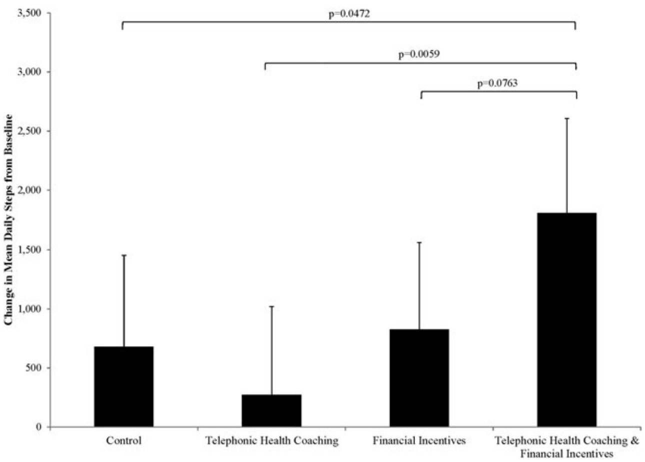

Figure 2. Average daily step count as measured by the Fitbit Zip among the 4 study arms adjusted for the baseline daily step count. The height of the bar illustrates the average daily step count 6 months following total knee replacement in the 4 study arms. The brackets indicate pairwise group comparisons with corresponding *P* values.

from all analyses due to missing baseline accelerometer data (0 valid wear days).

The results of this analysis were similar to the results of the main analyses. The mean increase in daily step count by 6 months compared to the pre-TKR levels for participants in the control arm was 544 (95% CI −165, 1,254); for participants in the THC arm, 368 (95% CI −346, 1,083); for participants in the FI arm, 924 (95% CI 195, 1,652); and for participants in THC + FI arm, 1,631 (95% CI 903, 2,358) (P = 0.0295 overall). The change in average daily steps over the 6-month period in THC + FI arm as compared to control arm was 1,086 (95% CI 76, 2,097). Three subjects (1%) withdrew from the study, and 14 subjects (7%) were discontinued for having undergone surgery requiring overnight hospitalization prior to 6 months. The results of the primary and sensitivity analyses did not change meaningfully after excluding these subjects.

DISCUSSION

In this randomized controlled trial of financial incentives and telephonic health coaching for increasing PA following TKR, we observed that a dual intervention consisting of both telephonic health coaching and financial incentives led to clinically relevant increases in step count (1,128 steps/day, effect size 0.4 SD) and PA (25 minutes/week, effect size 0.7 SD) compared to controls. Using distributional properties of the effect size (it is equivalent to Z score), an effect size of 0.4 SD indicates that two-thirds of the control arm would have step counts lower than the average in the THC + FI group, while an effect size of 0.7 SD would correspond to three-fourths of the control group having fewer minutes of PA than the average in the THC + FI group (44).

TKR alone does not lead to meaningful changes in PA (8,9). While surgery usually restores function and improves pain, previously sedentary individuals also require substantial behavior modification to change their postoperative PA. SPARKS made use of both internal and external motivation principles (coaching and incentives, respectively) to facilitate the development of appropriate PA habits in persons undergoing TKR. From our experience in this study, we would recommend that future interventions involving telephonic health coaching incorporate objective data captured in real time into the coaching sessions, rather than relying solely on participant self-report. Financial incentives, the other intervention examined, is a relatively new modality in the context of improving levels of PA. While the evidence for its efficacy is mixed (45,46), a recent robust randomized controlled trial demonstrated substantial effects of financial incentives in improving the health of sedentary persons (47). It appears that incentive amount, incentive structure, and demographic characteristics of the recipient population are all significantly related to intervention efficacy (48).

The results of SPARKS should be viewed in light of some limitations. We did not record the length of our attention control calls, and thus we are unable to make inferences about the effects of the duration of conversations. Our financial incentives were delivered with a lag period, blunting participants' situational associations between increasing PA and financial rewards. Additionally, only the participants receiving financial incentives were required to complete PA logs; PA diaries alone have been shown to increase step count (49). Similarly, participants were able to view their same-day step count on the Fitbit Zip, which could lead to an increased step count if participants used the Fitbit as a motivational device rather than simply a measurement device for the research team (50). Furthermore, reports published following the start of our trial have suggested that the Fitbit Zip may slightly overestimate individuals' activity levels (51). Given that all study participants wore devices of the same type, these issues are unlikely to affect our estimates of the efficacy of the interventions. Because Fitbit Zip measures steps and not intensity, we made the assumption that the Fitbit Zip would be less sensitive in detecting activity compared to medical-grade devices such as the ActiGraph and thus relaxed the generally accepted wear time criteria. The wear time thresholds have not been validated specifically for use with the Fitbit Zip. Only individuals with regular internet access were eligible to participate, which potentially limited the generalizability of our findings to TKR recipients with lower literacy levels or lower socioeconomic status; however, only 2% of otherwise eligible subjects who wanted to participate did not have internet access. As with all trials of modest sample size, fluctuations of just 2–3 subjects may lead to differences in participant baseline characteristics across arms. We examined these differences carefully and felt that they were within the boundaries of random fluctuation and that the 3 intervention arms did not meaningfully differ from the control arm. Finally, our intervention was limited in duration, spanning only 6 months following TKR.

The dual intervention consisting of financial incentives and telephonic health coaching during the first 6 months following TKR led to clinically meaningful increases in average step count for TKR recipients. With recent moves to shift health care financing from paying for volume to value (52), expenditure of resources to improve PA will become increasingly attractive, given that PA is among the most important drivers of quality of life in OA (53–57). Future studies should examine the impact of health coaching using real-time objective data in addition to financial incentives. Formal economic evaluation of these interventions should be undertaken to assess incorporating them into postoperative rehabilitation regimens in a cost-effective manner.

ACKNOWLEDGMENTS

The authors thank Drs. John Wright, Gregory Brick, Scott Martin, John Ready, and Thomas Thornhill, of Harvard Medical School and the Department of Orthopedic Surgery at Brigham and Women's Hospital, for referring their patients. The authors also thank Dr. Daniel Riddle of Virginia Commonwealth University for serving as the study safety monitor.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Losina had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study conception and design. Losina, Collins, Katz.
Acquisition of data. Deshpande, Smith, Michl, Usiskin, Klara, Winter, Selzer, Katz.
Analysis and interpretation of data. Losina, Collins, Yang.

REFERENCES

1. Hoffmann TC, Maher CG, Briffa T, Sherrington C, Bennell K, Alison J, et al. Prescribing exercise interventions for patients with chronic conditions. CMAJ 2016;188:510–8.
2. Schwingshackl L, Dias S, Hoffmann G. Impact of long-term lifestyle programmes on weight loss and cardiovascular risk factors in overweight/obese participants: a systematic review and network meta-analysis. Syst Rev 2014;3:130.
3. Khera AV, Emdin CA, Drake I, Natarajan P, Bick AG, Cook NR, et al. Genetic risk, adherence to a healthy lifestyle, and coronary disease. N Engl J Med 2016;375:2349–58.
4. McAlindon TE, Bannuru RR, Sullivan MC, Arden NK, Berenbaum F, Bierma-Zeinstra SM, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. Osteoarthritis Cartilage 2014;22:363–88.
5. Dunlop DD, Song J, Semanik PA, Chang RW, Sharma L, Bathon JM, et al. Objective physical activity measurement in the Osteoarthritis Initiative: are guidelines being met? Arthritis Rheum 2011;63:3372–82.
6. Losina E, Paltiel AD, Weinstein AM, Yelin E, Hunter DJ, Chen SP, et al. Lifetime medical costs of knee osteoarthritis management in the US: impact of extending indications for total knee arthroplasty. Arthritis Care Res (Hoboken) 2015; 67:203–15.
7. Beswick AD, Wylde V, Gooberman-Hill R, Blom A, Dieppe P. What proportion of patients report long-term pain after total hip or knee replacement for osteoarthritis? A systematic review of prospective studies in unselected patients. BMJ Open 2012;2:E000435.
8. Harding P, Holland AE, Delany C, Hinman RS. Do activity levels increase after total hip and knee arthroplasty? Clin Orthop Relat Res 2014;472:1502–11.
9. De Groot IB, Bussmann HJ, Stam HJ, Verhaar JA. Small increase of actual physical activity 6 months after total hip or knee arthroplasty. Clin Orthop Relat Res 2008;466:2201–8.
10. Brandes M, Ringling M, Winter C, Hillmann A, Rosenbaum D. Changes in physical activity and health-related quality of life during the first year after total knee arthroplasty. Arthritis Care Res (Hoboken) 2011;63:328–34.
11. Nilsdotter AK, Toksvig-Larsen S, Roos EM. Knee arthroplasty: are patients' expectations fulfilled? A prospective study of pain and function in 102 patients with 5-year follow-up. Acta Orthop 2009;80:55–61.
12. Goode AD, Reeves MM, Eakin EG. Telephone-delivered interventions for physical activity and dietary behavior change: an updated systematic review. Am J Prev Med 2012; 42:81–8.
13. Kivela K, Elo S, Kyngas H, Kaariainen M. The effects of health coaching on adult patients with chronic diseases: a systematic review. Patient Educ Couns 2014;97:147–57.
14. Sazlina SG, Browning C, Yasin S. Interventions to promote physical activity in older people with type 2 diabetes mellitus: a systematic review. Front Public Health 2013;1:71.
15. O'Donoghue T, Rabin M. The economics of immediate gratification. J Behav Decis Making 2000;13:233–50.
16. Loewenstein G, Prelec D. Anomalies in intertemporal choice: evidence and an interpretation. Q J Econ 1992;107:573–97.
17. Ainslie G. Derivation of "rational" economic behavior from hyperbolic discount curves. Am Econ Rev 1991;81:334–40.
18. Kahneman D, Tversky A. Choices, values, and frames. Am Psychol 1984;39:341–50.
19. Thomas S, Reading J, Shephard RJ. Revision of the Physical Activity Readiness Questionnaire (PAR-Q). Can J Sport Sci 1992;17:338–45.
20. Kirby KN, Petry NM. Heroin and cocaine abusers have higher discount rates for delayed rewards than alcoholics or non-drug-using controls. Addiction 2004;99:461–71.
21. Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS): validation and comparison to the WOMAC in total knee replacement. Health Qual Life Outcomes 2003;1:17.
22. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. Med Care 2005;43:203–20.
23. Nicholson N, Soane E, Fenton-O'Creevy M, Willman P. Personality and domain-specific risk taking. J Risk Res 2005;8:157–76.
24. Tang K, Beaton DE, Boonen A, Gignac MA, Bombardier C. Measures of work disability and productivity: Rheumatoid Arthritis Specific Work Productivity Survey (WPS-RA), Workplace Activity Limitations Scale (WALS), Work Instability Scale for Rheumatoid Arthritis (RA-WIS), Work Limitations Questionnaire (WLQ), and Work Productivity and Activity Impairment Questionnaire (WPAI) [abstract]. Arthritis Care Res (Hoboken) 2011;63 Suppl 11:S337.
25. Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. Pharmacoeconomics 1993;4:353–65.
26. Semanik P, Lee J, Manheim L, Dipietro L, Dunlop D, Chang RW. Relationship between accelerometer-based measures of physical activity and the Yale Physical Activity Survey in adults with arthritis. Arthritis Care Res (Hoboken) 2011;63: 1766–72.
27. Dipietro L, Caspersen CJ, Ostfeld AM, Nadel ER. A survey for assessing physical activity among older adults. Med Sci Sports Exerc 1993;25:628–42.
28. Collins JE, Rome BN, Daigle ME, Lerner V, Katz JN, Losina E. A comparison of patient-reported and measured range of motion in a cohort of total knee arthroplasty patients. J Arthroplasty 2014;29:1378–82.
29. Ware JE, Kosinski M, Dewey JE. How to score version 2 of the SF-36 health survey: standard and acute forms. QualityMetric; 2001.
30. Case MA, Burwick HA, Volpp KG, Patel MS. Accuracy of smartphone applications and wearable devices for tracking physical activity data [letter]. JAMA 2015;313:625–6.
31. Ferguson T, Rowlands AV, Olds T, Maher C. The validity of consumer-level activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. Int J Behav Nutr Phys Act 2015;12:42.
32. Tate LM. Temporal discounting rates and their relation to exercise behavior in older adults. Physiol Behav 2015;152: 295–9.
33. Miller W, Rollnick S. Motivational interviewing: preparing people for change. 2nd ed. New York: the Guilford Press; 2002.

34. Losina E, Collins JE, Daigle ME, Donnell-Fink LA, Prokopetz JJ, Strnad D, et al. The AViKA (Adding Value in Knee Arthroplasty) postoperative care navigation trial: rationale and design features. BMC Musculoskelet Disord 2013;14:290.
35. Volpp KG, John LK, Troxel AB, Norton L, Fassbender J, Loewenstein G. Financial incentive-based approaches for weight loss: a randomized trial. JAMA 2008;300:2631–7.
36. 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;40:181–8.
37. US Department of Health and Human Services. 2008 physical activity guidelines for Americans. URL: http://health.gov/paguidelines/guidelines/.
38. Tudor-Locke C, Sisson SB, Collova T, Lee SM, Swan PD. Pedometer-determined step count guidelines for classifying walking intensity in a young ostensibly healthy population. Can J Appl Physiol 2005;30:666–76.
39. Ayabe M, Aoki J, Kumahara H, Yoshimura E, Matono S, Tobina T, et al. Minute-by-minute stepping rate of daily physical activity in normal and overweight/obese adults. Obes Res Clin Pract 2011;5:E79–156.
40. Kroon FP, van der Burg LR, Buchbinder R, Osborne RH, Johnston RV, Pitt V. Self-management education programmes for osteoarthritis. Cochrane Database Syst Rev 2014;1:CD008963.
41. Volpp KG, Troxel AB, Pauly MV, Glick HA, Puig A, Asch DA, et al. A randomized, controlled trial of financial incentives for smoking cessation. N Engl J Med 2009;360:699–709.
42. Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB. Designing clinical research. Philadelphia (PA): Lippincott Williams & Wilkins; 2013.
43. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale (NJ): Routledge; 1988.
44. Coe R. It's the effect size, stupid. What effect size is and why it is important. In: Annual Conference of the British Educational Research Association. University of Exeter; 2002. URL: www.leeds.ac.uk/educol/documents/00002182.htm.
45. Molema CC, Wendel-Vos GC, Puijk L, Jensen JD, Schuit AJ, de Wit GA. A systematic review of financial incentives given in the healthcare setting: do they effectively improve physical activity levels? BMC Sports Sci Med Rehabil 2016;8:15.
46. Barte JC, Wendel-Vos GC. A systematic review of financial incentives for physical activity: the effects on physical activity and related outcomes. Behav Med 2017;43:19–90.
47. Patel MS, Asch DA, Rosin R, Small DS, Bellamy SL, Heuer J, et al. Framing financial incentives to increase physical activity among overweight and obese adults: a randomized, controlled trial. Ann Intern Med 2016;164:385–94.
48. Haff N, Patel MS, Lim R, Zhu J, Troxel AB, Asch DA, et al. The role of behavioral economic incentive design and demographic characteristics in financial incentive-based approaches to changing health behaviors: a meta-analysis. Am J Health Promotion 2015;29:314–23.
49. Bravata DM, Smith-Spangler C, Sundaram V, Gienger AL, Lin N, Lewis R, et al. Using pedometers to increase physical activity and improve health: a systematic review. JAMA 2007;298:2296–304.
50. Mansi S, Milosavljevic S, Baxter GD, Tumilty S, Hendrick P. A systematic review of studies using pedometers as an intervention for musculoskeletal diseases. BMC Musculoskelet Disord 2014;15:231.
51. Tully MA, McBride C, Heron L, Hunter RF. The validation of Fitbit Zip physical activity monitor as a measure of free-living physical activity. BMC Res Notes 2014;7:952.
52. Mechanic RE. Mandatory Medicare bundled payment: is it ready for prime time? N Engl J Med 2015;373:1291–3.
53. Lo GH, McAlindon TE, Hawker GA, Driban JB, Price LL, Song J, et al. Symptom assessment in knee osteoarthritis needs to account for physical activity level. Arthritis Rheumatol 2015; 67:2897–904.
54. McAlindon TE, Driban JB, Henrotin Y, Hunter DJ, Jiang GL, Skou ST, et al. OARSI clinical trials recommendations: design, conduct, and reporting of clinical trials for knee osteoarthritis. Osteoarthritis Cartilage 2015;23:747–60.
55. Messier SP, Callahan LF, Golightly YM, Keefe FJ. OARSI clinical trials recommendations: design and conduct of clinical trials of lifestyle diet and exercise interventions for osteoarthritis. Osteoarthritis Cartilage 2015;23:787–97.
56. Fitzgerald GK, Hinman RS, Zeni J Jr., Risberg MA, Snyder-Mackler L, Bennell KL. OARSI clinical trials recommendations: design and conduct of clinical trials of rehabilitation interventions for osteoarthritis. Osteoarthritis Cartilage 2015;23:803–14.
57. Dobson F, Hinman RS, Roos EM, Abbott JH, Stratford P, Davis AM, et al. OARSI recommended performance-based tests to assess physical function in people diagnosed with hip or knee osteoarthritis. Osteoarthritis Cartilage 2013;21: 1042–52.