

Original Paper

# Evaluating an Incentive-Based mHealth App for Physical Activity Promotion Using the Obesity-Related Behavioral Intervention Trial Model: Small Cohort Study

Babac Salmani<sup>1</sup>, PhD; Madison S Hiemstra<sup>1</sup>, MSc; Harry Prapavessis<sup>1</sup>, PhD; Leigh M Vanderloo<sup>2,3</sup>, PhD; Marc S Mitchell<sup>1</sup>, PhD

<sup>1</sup>School of Kinesiology, Faculty of Health Sciences, Western University, London, ON, Canada

<sup>2</sup>ParticipACTION, Toronto, ON, Canada

<sup>3</sup>School of Occupational Therapy, Faculty of Health Sciences, Western University, London, ON, Canada

## Corresponding Author:

Babac Salmani, PhD  
School of Kinesiology  
Faculty of Health Sciences, Western University  
1151 Richmond Street  
London, ON  
Canada  
Phone: 1 519-661-2111  
Email: [bsalmani@uwo.ca](mailto:bsalmani@uwo.ca)

## Abstract

**Background:** Physical inactivity remains a public health concern, with 42% (around 1 in 2) of women and 34% (around 1 in 3) of men in the United Kingdom, for example, failing to meet moderate-to-vigorous physical activity guidelines. To promote physical activity (PA) at scale, smartphone-based mHealth (mobile health) software apps offer a promising solution.

**Objective:** This study aims to evaluate the feasibility of implementing an mHealth app offering very small (“micro”) financial incentives for PA in Leeds, United Kingdom.

**Methods:** A 5-week single-arm proof-of-concept study was conducted with rolling recruitment among Caterpillar Health app users between September 12 and December 12, 2022 (Obesity-Related Behavioral Intervention Trial model, phase IIa). Users earned microincentives in the form of “points,” redeemable for consumer rewards (eg, movie tickets and gym passes), for meeting personalized daily step goals (US \$0.13 per goal achieved; set using data from a 5-day baseline) and completing educational quizzes (US \$0.33 per quiz). Descriptive statistics assessed feasibility outcomes (ie, reach, recruitment, retention, engagement, and acceptability) and preliminary effectiveness. Paired-samples *t* tests ( $P < .05$ ) examined changes in weekly mean daily step count (from baseline) and step goal achievement over 5 weeks.

**Results:** Of 285 app downloads, 46 users consented to participate (recruitment rate: 16.1%). Participants (mean age: 39.9, SD 11.1 y; 71.1%, 33/46 woman) had a baseline step count of 5598 (SD 2664) steps/day. A total of 25 participants remained engaged (ie, completed at least 1 quiz) at study week 5 (retention rate: 54.3%). Acceptability was high, with 75% of respondents (12/16) indicating they would recommend the app. Weekly mean daily step count did not significantly increase from baseline (mean difference 317, SD 2273,  $P = .53$ ). Weekly daily step goal achievement rate (%) decreased from study week 1 to 5 (−23.23, SD 22.85,  $P = .02$ ).

**Conclusions:** Despite lower-than-expected recruitment and no statistically significant PA increase, relatively high engagement and acceptability suggest future pilot testing (Obesity-Related Behavioral Intervention Trial model, phase IIb) of a refined intervention (eg, wider selection of loyalty reward partners) and modified study protocol (eg, simplified consent process) is warranted.

**Trial Registration:** ClinicalTrials.gov NCT05294692; <https://clinicaltrials.gov/study/NCT05294692>

*JMIR Form Res* 2026;10:e85484; doi: [10.2196/85484](https://doi.org/10.2196/85484)

**Keywords:** mHealth; digital health; financial incentives; health behavior change; physical activity; behavioral economics

## Introduction

Physical inactivity remains a public health concern, with 42% (around 1 in 2) of women and 34% (around 1 in 3) of men in the United Kingdom, for example, failing to meet moderate-to-vigorous physical activity guidelines [1]. To promote physical activity (PA) at scale, smartphone-based mHealth (mobile health) software apps offer a promising solution [2]. As global smartphone ownership continues to rise (ie, 85% in 2024 vs 66% in 2018 worldwide [3], with 87% of UK adults owning a smartphone in 2023) [4], mHealth apps may be one way of promoting population-level PA [5,6]. Despite their potential, low mHealth app engagement (eg, frequency and depth of use; a common precondition for mHealth app effectiveness) leading to little or no effect is typical [7]. Grounding mHealth app intervention in behavior change theory (eg, behavioral economics) may boost engagement and effectiveness [7-9]. For example, a systematic review of 35 randomized controlled trials (RCTs) by Boonmanunt et al [8] found that interventions informed by theory outperformed those that were not.

Behavioral economics, a branch of economics complemented by insights from psychology, has shown promise in boosting mHealth app engagement [10]. One of the central tenets in behavioral economics is that of “decision bias.” “Decision bias” refers to the systematic ways in which human decisions deviate from what would be considered rational or logical [11]. For example, “present bias” refers to the tendency for individuals to place disproportionate emphasis on the present costs and benefits of behaviors in decision-making while at the same time discounting future ones [11]. In the context of PA, for example, the immediate costs experienced with a bout of exercise (eg, discomfort) often outweigh future benefits (eg, lower cancer risk) in daily decision-making [12]. Financial incentives (FI), monetary rewards contingent on health behavior or outcome goal achievement, may encourage individuals to engage in more PA today, for instance, given the added short-term benefit of rewards [12]. There is evidence that FI-for-PA interventions can be effective [13,14]. For example, in a review of RCTs examining PA incentives, Mitchell et al [14] found that FI, worth on average about US \$1.50 per day, produced short-term (ie, <6 mo), long-term (ie, ≥6 mo), and follow-up (ie, after incentive withdrawal) effects. While promising, these FI magnitudes may be cost-prohibitive when implemented at a population level, or longer-term (ie, more than 6 months, a theoretical threshold of behavior maintenance) [15]. This may be particularly true within the constraints of limited public health budgets [16]. For example, offering £1.00 (ie, US \$1.33/d) per day to just 1% of the UK adult population (ie, 683,500 people) would cost £249.5 million (US \$334.8 million) annually (ie, about 4.82% of the 2025 £3.85 billion [US \$5.17 billion] public health budget) [17].

One way of controlling FI costs while maintaining their effects may be to limit the size of rewards, all the while maintaining integrity to behavioral economics—that is, offering FI close in time to the desired behavior or outcome to best leverage “present bias.” In the 2020 Mitchell et al

[14] review, several RCTs offering very small (“micro”) FI increased PA (eg, US \$0.09-US \$0.53/d) [18-20]. For example, Rohde et al [18] offered microincentives (ie, US \$0.09-US \$0.15/d) conditional on participants’ weekly gym attendance (ie, FI per day). These small cash rewards, given to both regular and irregular gym visitors, led to increased attendance in both the short- (ie, <12 wk) and long-term (ie, ≥12 wk) compared to a control group receiving unconditional FI. Much of the evidence supporting microincentive impact is derived from RCTs with limited external validity, however. Few have examined their impact in real-world public health settings [21-23]. Knowing more about microincentives outside of controlled experimental settings may enhance their impact in the public sphere.

One example of microincentives for PA deployed in the real world is Carrot Rewards (2015-2019) [24], a Canadian mHealth app that offered microincentives for daily step goal achievement (eg, consumer loyalty points worth about US \$0.03/day [ie, equivalent to 1 loyalty reward point] for movies, groceries, gas, etc). On one hand, the app saw high uptake and engagement, attracting more than 1.3 million downloads and 500,000 monthly active users and produced short- (ie, <6 mo) and long-term (ie, ≥6 mo) PA improvements [24]. On the other hand, higher than anticipated engagement resulted in large-scale daily step goal reward payouts (ie, totaling more than US \$3 million over 4 years). This great cost, combined with reliance on government funds to pay for the microincentives [25], undermined the app’s scalability or sustainability potential and ultimately led to its discontinuation in 2019 [24]. Given the strong uptake, high engagement, and long-term effects seen with Carrot [16, 24], efforts to replicate their microincentive approach, with some modification (eg, wean users off daily step goal rewards after a year), in different contexts are warranted. Specifically, examining how digitally delivered microincentives for PA may perform in a completely different country (ie, United Kingdom) with different partners (eg, loyalty point programs, government agencies, and nongovernmental organizations) and different users (eg, who may not respond to points or incentives the same way) [26] may yield valuable insights.

The purpose of this small 5-week proof-of-concept study, therefore, is to evaluate the feasibility of offering microincentives for PA in the United Kingdom via a commercial mHealth app. The primary indicators for this study are feasibility-related outcomes (eg, uptake, engagement, and acceptability), rather than statistically significant changes in PA. This study represents an important first step in a multiphased project aiming to develop, refine, and evaluate a promising (ie, more scalable and sustainable) FI-for-PA intervention in the United Kingdom.

## Methods

### Study Setting and Design

This was a small 5-week single-arm proof-of-concept study with rolling recruitment between September 12 and December 12, 2022. This research was conducted in partnership

with Caterpillar Health, an mHealth app designed to support PA engagement through personalized daily step goals, evidence-supported educational content, and microincentives in Leeds, United Kingdom. This preparatory study was guided by the Obesity-Related Behavioral Intervention Trial (ORBIT) model [27], a structured framework that supports the systematic development and evaluation of behavioral interventions. The ORBIT model outlines staged phases, from early feasibility (phase I-II) to large-scale effectiveness testing (phase III-IV). This study represents phase IIa (ie, proof-of-concept), which emphasizes feasibility assessment as well as identifying opportunities to refine the intervention and future study design. To ensure the intervention and study design incorporated principles of equity, diversity, inclusion, and accessibility, the World Health Organization's health equity toolkit [28] was used. This included appropriately describing categories of disadvantage (eg, gender, household income, and ethnicity), avoiding overly technical health language in both the app interface and recruitment materials, using inclusive imagery within the app, ensuring compatibility across different mobile operating systems (eg, Android [Google LLC] and iOS [Apple Inc]), and tailoring the approach to Leeds, United Kingdom (ie, a community with a relatively high prevalence of overweight and obese adults [63.9%]) [29] to ensure sensitivity to local health inequities. In line with the World Health Organization toolkit's emphasis on participatory design, feedback from app users within the community was sought during this study to identify potential barriers and ensure cultural relevance. In addition, subgroup analyses were conducted to examine the impact of the interventions among equity-deserving groups (eg, individuals with low income, chronic conditions, or low baseline PA) to assess whether the strategies were equitably effective across diverse populations.

To promote the app, a paid advertising campaign on social media platforms (ie, Facebook [Meta], Instagram [Instagram from Meta], and Twitter [subsequently rebranded X; X Corp]) began on August 1, 2022 (ie, outside of this study's period) to raise awareness of the app leading up to its launch on September 12, 2022. To further support app launch, Caterpillar collaborated with Leeds City Council to have emails sent to City employees, as well as to local KPMG (Klynveld Peat Marwick Goerdeler) employees and Hussle Gym members living in Leeds, United Kingdom, on August 8, 2022. Participation required several sequential steps, including app download and account registration. Informed consent to participate in this study was sought on 2 separate occasions. First, to authorize the collection of participants' sociodemographics and app engagement data (starting September 12, 2022) through an in-app consent form for study participation, and then again separately (starting

September 30, 2022) to authorize the collection of device-assessed daily step count data over a 5-day baseline followed by a 5-week intervention period. This aligned with the app's "steps" feature launch on September 30, 2022, and captured participants' daily step counts in the approximately 6 weeks following the second consent. Sociodemographic and health characteristics were collected through a survey embedded in the app ([Multimedia Appendix 1](#)). This trial was registered at ClinicalTrials.gov (NCT05294692). The STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement checklist ([Checklist 1](#)) [30] for cohort studies is provided.

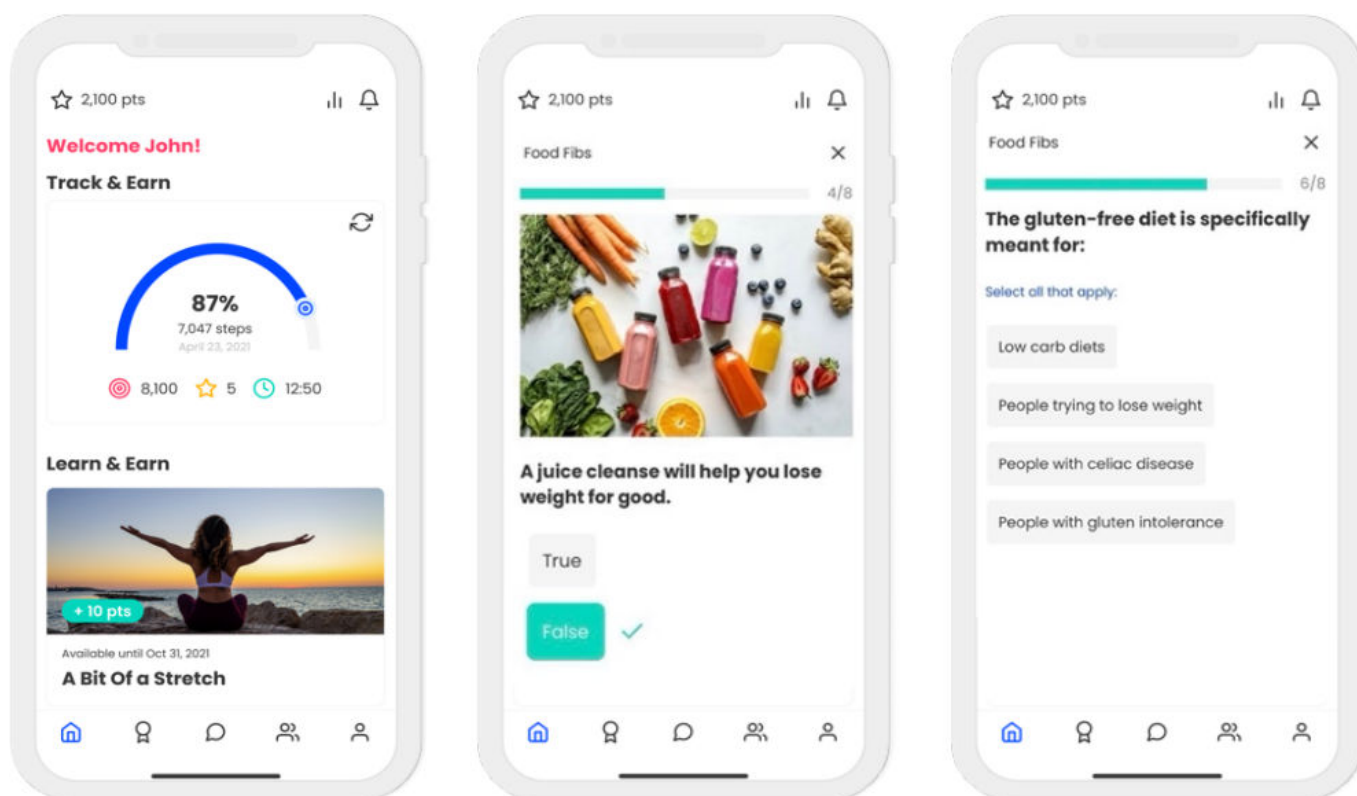
## **Ethical Considerations**

This trial was approved by Western University Health Sciences' Research Ethics Board in March 2024 (Project ID: 120615). All participants provided informed consent before participation. Participant privacy and confidentiality were maintained through the use of deidentified data, secure data storage systems, and restricted access to study records. Participants were not compensated for study participation; however, as part of the intervention, they were eligible to earn FI contingent on achieving daily step goals. Incentives were provided in British pounds (£0.10), equivalent to approximately US \$0.13, based on a conversion rate of £1=US \$1.34 at the time of this study.

## **Intervention**

Caterpillar's cornerstone feature, "personalized challenges," offered microincentives for personalized daily step goal achievement ([Figure 1](#)). Users' initial daily step goals were calculated as the median step count from the 5-day (ie, no goal and no microincentive) baseline period, where users were instructed to "wear their device" as much as possible. After this baseline, users could earn microincentives in the form of points ("incentives and rewards"; equivalent to US \$0.13/d) for meeting their daily goals (ie, calculated as the median step count from the previous 7-day week). These points could be redeemed for consumer goods (ie, movie tickets from Vue Cinemas and gym passes from Hussle Gym). Users could also opt into weekly "step challenges" (ie, complete 5 out of 7 daily step goals) for additional points (US \$0.33). Finally, users could also complete 4 educational quizzes per week ("health education and microlearning"; US \$0.33/quiz). The educational quizzes primarily covered UK healthy eating and PA guideline content (in line with the National Health Service, "live well" guidelines). The app's key features, along with their alignment to relevant behavioral economic constructs, are described in full in [Multimedia Appendix 2](#).

**Figure 1.** The Caterpillar Health mobile app allowed participants to monitor their PA, view progress toward goals, and interact with educational content designed to support engagement with the platform. PA: physical activity.



## Outcomes

Several feasibility outcomes are reported, including (1) reach (ie, number of emails sent during app promotion), (2) recruitment (ie, number of app downloads and number of users providing initial informed consent) and retention (ie, proportion of participants completing an educational quiz in study week 5), (3) app engagement (ie, proportion of total possible educational quizzes completed during this study's period), and (4) intervention acceptability (ie, exit survey responses informed by the validated Mobile Application Rating Scale [31,32]; Multimedia Appendix 1) data were collected. To examine the potential impact of the app on PA, device-assessed (5) baseline and weekly mean daily step count (ie, 1000 to 40,000 steps/day considered within an acceptable range) was collected over a roughly 6-week period (ie, 5 days plus 5 weeks following "steps" feature launch), while (6) daily step goal achievement (ie, proportion of days [out of 7] a user met or exceeded their daily step goals each week) was assessed over the 5-week intervention period.

## Theoretical Underpinning

The behavior change wheel (BCW) [33], an amalgamation of 19 behavior change theories, provides a structured and theory-driven framework for designing and evaluating behavioral interventions aimed at both short- and long-term behavior change. The BCW consists of 3 layers. At its core, the BCW defines behavior as an interaction between one's capability (eg, knowledge and skills), opportunity (eg, access to facilities and social support), and motivation (eg, goal setting and rewards) to engage in a particular health behavior. This is referred to as the capability, opportunity,

motivation–behavior (COM-B) model. The integration of the COM-B model with microincentives may help tackle barriers to PA (eg, lack of knowledge and low engagement) and enhance the effectiveness of future interventions to sustain behavior over time. To date, several commercially available PA apps (ie, Sweatcoin, Carrot Rewards, StepBet, etc) have incorporated behavior change theories (ie, self-determination theory) within their app design, alongside FI mechanisms (ie, gamified points and lotteries) to boost PA and drive user engagement. However, further research may be needed to test the effectiveness of microincentives in real-world settings. During the design phase of the Caterpillar Health app, the research team used the COM-B model as the theoretical foundation of the app to identify and address barriers and enablers to increase daily step counts and app engagement. Specifically, the model guided the identification of key influencing factors (ie, lack of motivation and time constraints) and facilitators (ie, social support and user-friendly interfaces). These insights were then used to inform how the app's features could be designed to best elicit behavior change. The categories evaluate the intervention derived from the behavior change technique taxonomy, which systematically classifies 93 distinct techniques into 16 clusters used to design and evaluate behavior change interventions [34,35]. The app's key features are described according to the COM-B model in Multimedia Appendix 2.

## Statistical Analysis

Statistical analysis was performed using SPSS (version 28.0.1.0; IBM Corp). Descriptive statistics were used to assess reach, recruitment, and retention rates, app

engagement, and intervention acceptability. To assess preliminary effectiveness, paired-samples *t* tests were used to examine change in weekly mean daily step count from the 5-day baseline (vs study week 5) among the subsample of users downloading the app on or after September 30, 2022 (with “steps” feature included). Exploratory analyses also examined change in weekly mean daily step count by PA level (ie, physically active:  $\geq 5000$  steps/d at baseline; physically inactive:  $< 5000$  steps/d at baseline) [36], engagement (ie, high engagement:  $\geq 50\%$  of total quizzes completed [ $\geq 12$  of 24 quizzes]; low engagement:  $< 50\%$  of quizzes completed [ $< 12$  of 24 quizzes]), and chronic condition status (ie, at least 1 chronic condition vs no chronic conditions). Given the small sample size in each of these subgroups (Figure 2), the mean difference (MD) in weekly mean daily step counts and daily step goal achievement was reported descriptively. As this was a proof-of-concept study, a priori sample size calculation to detect statistical significance was not conducted. The results, however, will be used to inform sample size calculations for future pilot work. Nonetheless, statistical significance was measured as  $P < .05$ , and Cohen *d* for local effect sizes of weekly mean daily step counts were calculated, where  $d = 0.2$  indicates a small effect,  $d = 0.5$  a medium effect, and  $d = 0.8$  a large effect.

## Results

### Sample Characteristics

A total of 285 users downloaded the app, and 228 completed app registration. Of these, 46 users provided initial informed consent (Table 1). A subgroup of users also provided consent authorizing the collection of device-assessed daily step count data ( $n = 22$ ). Among the total sample, participants were primarily woman (71.1%, 32/45) with an average age of 39.9 years and a baseline daily step count of 5598 steps/day. Additionally, 65.7% (23/35) self-reported one or more chronic condition diagnoses, and 25% (9/36) reported at least one disability or impairment. Notably, a greater proportion of participants reported the use of a fitness tracker (65.9%, 27/41) and lower household income (18.9%, 7/37) compared to the general UK population (35% and 13%, respectively). These demographic characteristics may have influenced engagement and responsiveness to app-based microincentives, potentially limiting the generalizability of findings to broader populations (eg, more balanced gender distribution, lower fitness tracker use, and different health profiles).

**Table 1.** Baseline characteristics of the total sample, subgroup, and general UK population. Chronic condition defined as a diagnosis by a health professional with symptoms expected to last, or have already lasted 6 months or more, “total” sample includes participants who provided initial informed consent for data collection, “subgroup” sample includes participants who provided a separate informed consent for step count tracking. Household income of US \$27,156/year is equivalent to the UK’s cutoff for low household income (£21,000/year).

Characteristics	Total (N=46)	Subgroup (n=22)	UK population [36,37] (n=67,749,091)
Age (years)	39.9 (11.1) <sup>a</sup>	43.2 (10.5) <sup>a</sup>	40.7 <sup>b</sup>
Gender (woman) <sup>c</sup> , % (n/N)	71.1 (32/45)	86.4 (19/22)	51.0
Ethnicity (% White) <sup>d</sup>	80.5 (33/41)	95.5 (21/22)	81.7
BMI (kg/m <sup>2</sup> )	27.8	29.5	27.4
Chronic condition (one or more), % (n/N)	65.7 (23/35)	65.0 (13/20)	38.4
Diagnosed disability, % (n/N)	25.0 (9/36)	19.1 (4/21)	24.0
Housing location (urban), % (n/N)	89.2 (33/37)	95.5 (21/22)	84.2
Operating system (iOS), % (n/N)	62.2 (23/37)	95.5 (21/22)	50.8
Fitness tracker (yes), % (n/N)	65.9 (27/41)	77.3 (17/22)	35.0
Household income (below US \$27,156/year), % (n/N)	18.9 (7/37)	5.3 (1/19)	13.0
Steps per day (baseline mean, SD)	5598 (2664)	5547 (2619)	5444 (N/A) <sup>e</sup>

<sup>a</sup>Mean (SD).

<sup>b</sup>As the UK population data were obtained from the Office for National Statistics (2021 Census), only the median age in years is available. The IQR is not reported in this source and therefore cannot be included.

<sup>c</sup>Gender was self-reported by participants based on their gender identity (e.g., man, woman).

<sup>d</sup>Participants selected ‘Caucasian’ in the original questionnaire; this has been reported as ‘White’ in the manuscript to reflect current terminology.

<sup>e</sup>N/A: not applicable.

### Main Findings

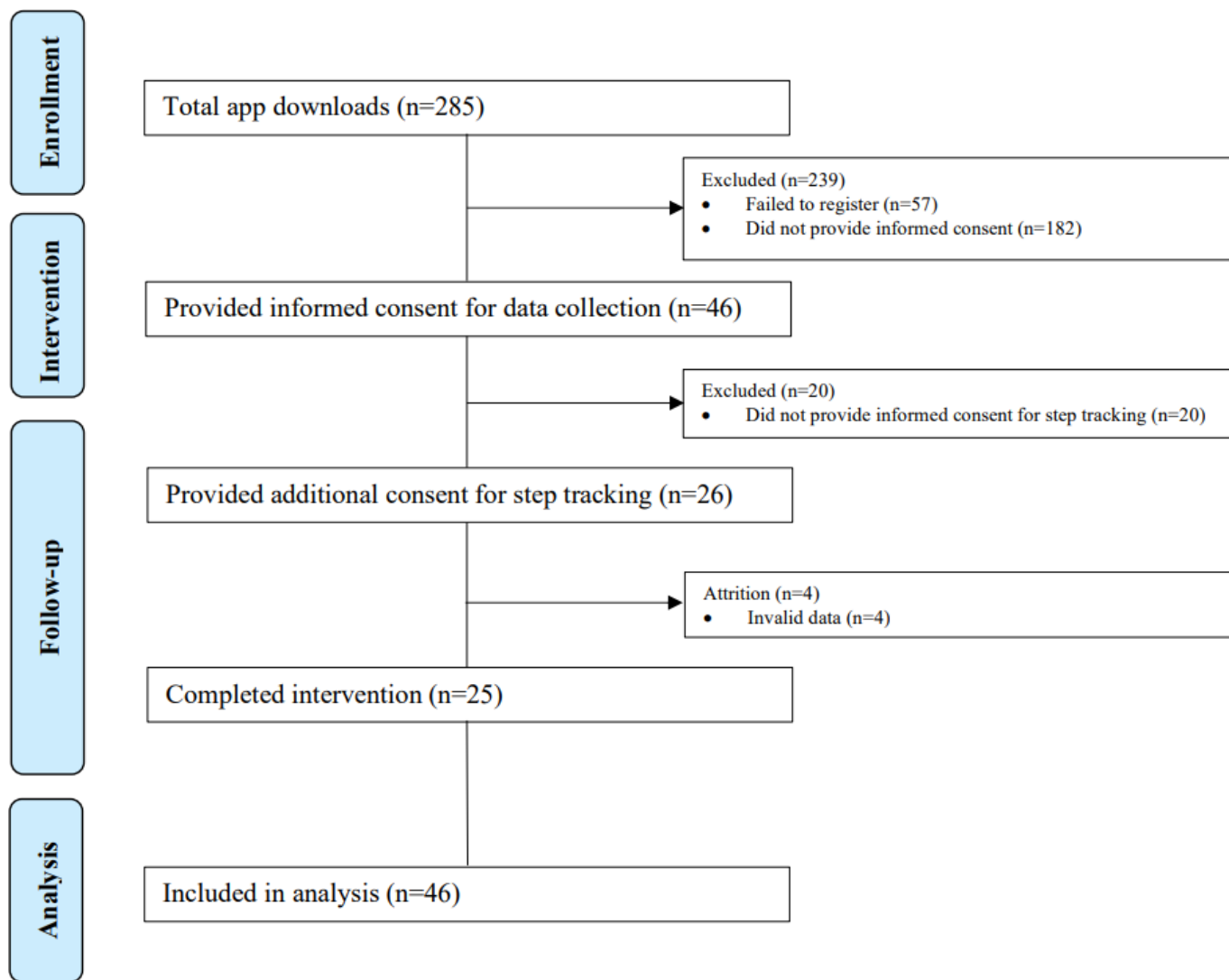
In total, over 15,000 and 3000 emails were sent to Leeds City Council and KPMG employees, respectively, while more than 7000 emails targeted Huddle Gym members. The number of social media impressions was not recorded. During this study’s period, Caterpillar recorded 285 app downloads, of which 46 (16%) provided informed consent and enrolled in

this study. Retention at the end of the 5-week intervention was 54.3% (25/46 enrolled participants; Figure 2). Among participants providing initial consent, 95.6% ( $n = 44$ ) completed at least one educational quiz, and 39.1% ( $n = 18$ ) met the criterion for high engagement, defined as completion of  $\geq 50\%$  of available quizzes ( $\geq 12$  of 24 quizzes). On average, participants completed 8.1 quizzes over the 5-week intervention period (out of 24 quizzes offered). Intervention

acceptability was high, with most participants completing the exit survey reporting “good” to “excellent” app aesthetics (15/16, 93.7%), having learned new health information (14/16, 87.5%), and, notably, they would recommend the app to others (12/16, 75%). Constructive feedback focused on

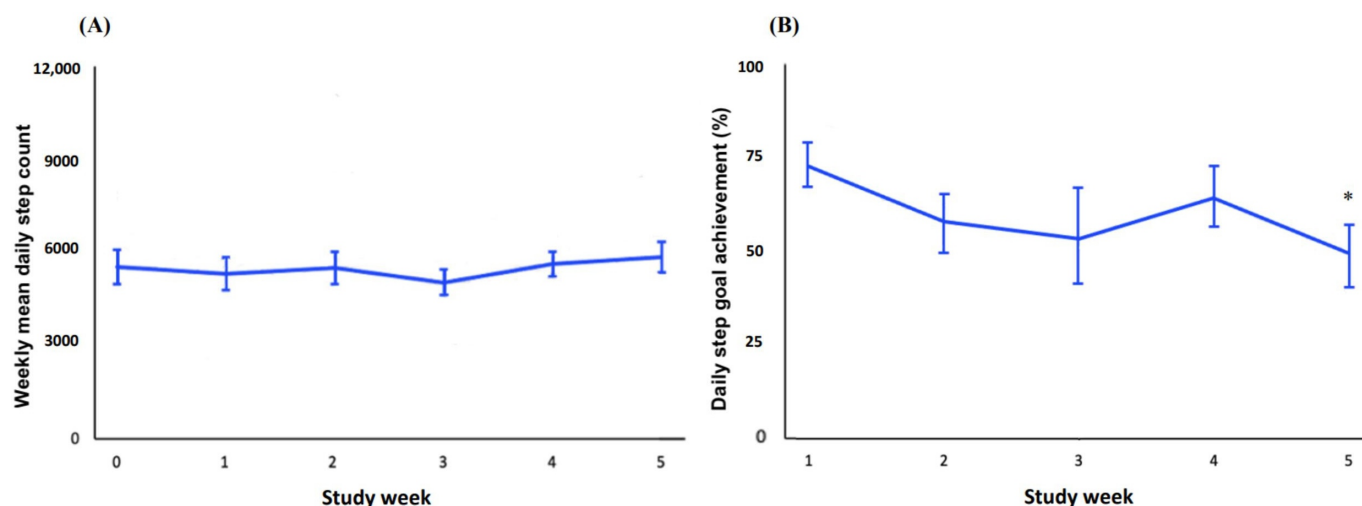
the lack of customizable features (eg, user input and sharing options; 2/16, 12.5%) and dissatisfaction with daily step goal presentation (eg, not visually engaging or personalized to the user’s preferences; 2/15, 13.3%).

**Figure 2.** Flow chart of participants in a 6-week study with rolling recruitment between September 12 and December 12, 2022, using the Caterpillar Health app. Failed to register refers to users who attempted to sign up for the app but did not complete the registration process. Did not provide informed consent includes users who accessed the platform but did not complete the consent process required for participation in this study. Invalid data includes users with less than 4 days in a week with step counts in the acceptable range (ie, 1000 to 40,000) during the 5-week physical activity assessment period (step count data were included for the 22 users who provided consent for step tracking). Completed intervention represents users who completed at least 1 quiz in study week 5.



Weekly mean daily step count did not significantly increase from baseline (MD 317, SD 2273,  $P=.53$ ; [Figure 3](#)). Moreover, weekly daily step goal achievement rate dropped over the 5-week period (MD  $-23.23$ , SD 22.85,  $P=.02$ ). Among inactive participants (ie,  $<5000$  steps at baseline), the weekly mean daily step count at baseline and study week 5 was 3509 (SD 1024) and 4656 (SD 2264), respectively ([Multimedia Appendix 3](#)). In contrast, active participants accumulated 7585 (SD 2053) steps/day at baseline and 7073 (SD 1753) at study week 5. Participants with at least one chronic condition had a mean daily step count of 5104 (SD

2533) at baseline and 5734 (SD 2434) at study week 5, whereas those without chronic conditions accumulated 6279 (SD 2932) and 5853 (SD 2379), respectively. Weekly step goal achievement rates were 74.2 (SD 18.6) at week 1 and 62.8 (SD 23.9) at study week 5 among inactive participants, and 76.2 (SD 16.5) and 33.3 (SD 8.3), respectively, among active participants. Those with chronic conditions reported 74.3 (SD 18.6) at week 1 and 54.2 (SD 27.5) at study week 5; those without chronic conditions reported 76.2 (SD 16.5) and 47.6 (SD 21.8), respectively.

**Figure 3.** Weekly mean (A) daily step count and (B) daily step goal achievement with standard errors.

## Discussion

### Principal Results

This single-arm, proof-of-concept ORBIT model phase IIa study assessed the feasibility of implementing Caterpillar, a commercial mHealth app offering microincentives for PA in Leeds, United Kingdom. Over 25,000 emails were sent on August 8, 2022, to raise awareness of the app launch on September 12, 2022. The app was downloaded 285 times during this study's period, with approximately 15% (46/285) of users providing initial consent to participate in this study. Over the 5-week intervention, about 40% (18/46) of users were classified as very engaged (ie, completed  $\geq 50\%$  of quizzes [ $\geq 12$  of 24 quizzes]). Intervention acceptability was high, with 75% (12/16) indicating they would recommend the app to others. Although Caterpillar did not appear to significantly increase PA, daily step goal achievement was fairly high in study week 1 (about 75%), along with exploratory subgroup trends (eg, for inactive adults), suggesting potential for impact. This finding should be interpreted in the context of a sample with relatively high baseline activity levels, which may have limited the scope for short-term, measurable increases in PA within this low-intensity, microincentive intervention. The discrepancy between high engagement and limited PA change may reflect several factors, including the short 5-week intervention duration, constraints in the incentive structure (eg, limited reward options), and potential limitations in how PA was measured (eg, reliance on daily step counts only). As this is a feasibility study, the emphasis is on effect sizes and feasibility outcomes (eg, engagement and acceptability) rather than on statistical significance; *P* values are reported for reference only. Taken together, these findings suggest the delivery of PA microincentives via Caterpillar in Leeds is feasible, with PA promotion potential given relatively high engagement and favorable subgroup trends. Pilot testing to assess the preliminary effectiveness of a refined intervention and study protocol is warranted (ie, phase IIb).

Several opportunities to improve the intervention and study protocol were identified (Table 2). For example, under

“rewards and threats” in Table 2, only 2 corporations (ie, Vue Cinemas and Hussle Gym) partnered with Caterpillar to deliver consumer rewards that could be redeemed through the app, potentially limiting the app's appeal. Partnerships with a wider array of corporations to provide more rewards options may help boost uptake in the future, as suggested by studies demonstrating that perceived reward value may help drive mHealth app engagement (eg, Are rewards meaningful to the target audience?) [38,39]. Under “feedback and monitoring” in Table 2, the app targeted any-intensity daily step count to, in part, promote participant autonomy and the potential for sustained change. Future iterations could broaden the intervention to promote a wider range of PA intensities (eg, moderate-to-vigorous physical activity and high-intensity interval training) and behaviors (eg, cycling, resistance training, and sedentary behavior reduction). In addition, under “associations” in Table 2, participant feedback highlighted limited opportunities for customization (eg, setting routine-based reminders and personalized push notifications), which will directly inform future iterations of Caterpillar to enhance usability and sustained engagement. From a study protocol perspective, several enhancements could strengthen future iterations of this intervention. First, more active recruitment strategies (eg, in-person presentations and text message invitations) could be used in place of the more passive approaches used here (eg, company-wide emails and social media campaigns) [40,41]. Second, the current goal-setting approach, which relies on the user's median daily step count from a 5-day baseline and a 7-day window for each subsequent week, may be suboptimal, particularly given the observed decline in step goal achievement over time. Short baselines may not fully capture typical variability in daily step patterns, potentially resulting in goals that are too easy or too difficult and limiting motivation [33]. Future trials could improve goals by using a longer baseline period (eg, 7 days), rolling averages across multiple weeks (eg, median step count over the past 2 or 3 weeks), and goals that account for variability in step count (eg, account for both median step count and SDs), potentially offering better daily goals [42,43]. Finally, the current single-arm design limits the ability to draw causal inferences regarding the intervention's

effectiveness. Later phases of research could include a pilot intervention effectiveness and to inform the design of future RCT (ie, ORBIT phase IIb) to allow for preliminary testing of fully powered trials [27].

**Table 2.** Intervention and study protocol areas for improvement and recommendations. Valid data includes users with daily step counts in an acceptable range (ie, 1000 to 40,000). The categories evaluating the intervention derive from the behavior change technique taxonomy [34,35].

	Areas for improvement	Evidence-informed recommendations
<b>Intervention</b>		
Associations	<ul style="list-style-type: none"> <li>Few prompts or contextual cues to reinforce behavior</li> </ul>	<ul style="list-style-type: none"> <li>Introduce customizable push notifications, routine-based reminders, prompts (eg, after meals), and environmental cues (eg, “take the stairs instead of the elevator”) [44]</li> </ul>
Comparison of behavior	<ul style="list-style-type: none"> <li>No social comparison or benchmarking</li> </ul>	<ul style="list-style-type: none"> <li>Add anonymized leaderboards, user averages, or percentile rankings to show how user performance compares with peers [45]</li> </ul>
Feedback and monitoring	<ul style="list-style-type: none"> <li>Only step count tracked; PA<sup>a</sup> intensity, duration, and type were not captured</li> </ul>	<ul style="list-style-type: none"> <li>Expand intervention to include other PA types (eg, cycling, resistance training, and sedentary behavior reduction) [46]</li> <li>Display personalized insights (eg, comparative analysis of PA behavior to users’ previous week) [47]</li> </ul>
Repetition and substitution	<ul style="list-style-type: none"> <li>Limited support for habit formation or alternative PA routines</li> </ul>	<ul style="list-style-type: none"> <li>Include streak tracking and options to substitute missed step goals with alternative PA tasks (eg, cycling) [48]</li> </ul>
Rewards and threats	<ul style="list-style-type: none"> <li>Limited consumer reward options (n=2)</li> <li>Boost microincentive impact with promising incentive designs</li> </ul>	<ul style="list-style-type: none"> <li>Offer additional reward options (eg, grocery, transit, and entertainment) [49]</li> <li>Test lost-framed incentives [49]</li> </ul>
Shaping knowledge	<ul style="list-style-type: none"> <li>Educational content delivered only through static quizzes and educational content</li> </ul>	<ul style="list-style-type: none"> <li>Incorporate interactive learning modules, infographics, video explainers, and progressive unlocking of content based on user performance [50]</li> </ul>
Social support	<ul style="list-style-type: none"> <li>No mechanisms to connect users or enable peer interaction</li> </ul>	<ul style="list-style-type: none"> <li>Add features such as group challenges, friend invitations, discussion boards, or accountability buddies [51]</li> </ul>
<b>Study protocol</b>		
Study setting	<ul style="list-style-type: none"> <li>Limited generalizability beyond the Leeds, United Kingdom context</li> </ul>	<ul style="list-style-type: none"> <li>Replicate procedures in diverse geographies (eg, urban and rural North American cities) [47]</li> </ul>
Study design	<ul style="list-style-type: none"> <li>Single-arm design limits causal inference (ie, no comparator group to isolate app effects)</li> </ul>	<ul style="list-style-type: none"> <li>Consider (pilot) RCT<sup>b</sup> design, or incorporate a waitlist condition, in future phases for improved causal inference [27]</li> </ul>
Recruitment	<ul style="list-style-type: none"> <li>Reliance on passive outreach strategies (eg, social media and company-wide email)</li> <li>Lack of tailored messaging across channels</li> <li>Rolling recruitment may lead to variability in exposure time and seasonal bias</li> <li>Staggered consent process may have confused participants, limited recruitment rate</li> </ul>	<ul style="list-style-type: none"> <li>Add more active recruitment, such as at community health events or workplace briefings, to reach underserved groups (eg, low-income adults) [52]</li> <li>Customize messaging based on sociodemographics (eg, younger adults via Instagram and older adults via email with testimonials) [52]</li> <li>Standardize recruitment windows or stratify analysis by recruitment timing to reduce confounding [53]</li> <li>Combine consents or replace with an integrated consent process [54]</li> </ul>
Sample	<ul style="list-style-type: none"> <li>Limited sample diversity (eg, 71% [32/45] woman, 81% [33/41] White<sup>c</sup>)</li> <li>High prevalence of fitness tracker use (65.9% [27/41]) may limit broad applicability</li> <li>Small sample size (n=46; n=22)</li> </ul>	<ul style="list-style-type: none"> <li>Consider oversampling underrepresented groups or tailoring eligibility to target populations (eg, low-income or visible minority groups) using stratified sampling (ie, randomly sample participants from distinct strata [eg, by age group and gender]) [55]</li> <li>Prioritize strategies to engage individuals who do not have access to extra digital tools at baseline (eg, fitness trackers) [55]</li> </ul>
Outcome assessment	<ul style="list-style-type: none"> <li>Calculation of users’ initial daily step goals (median step count from the 5-day baseline) may not be ideal</li> <li>No data collected beyond 6 weeks of app use</li> <li>Sole reliance on quiz completion for engagement measurement</li> </ul>	<ul style="list-style-type: none"> <li>Use a full 7-day baseline period using only valid data to better represent habitual PA patterns and account for weekday or weekend differences [56]</li> <li>Add postintervention follow-up to assess maintenance or motivation [57]</li> <li>Track additional engagement metrics (eg, number of app opens, time spent per open, and reward redemption rates) [58]</li> </ul>

	Areas for improvement	Evidence-informed recommendations
Analysis	<ul style="list-style-type: none"> <li>Paired sample <i>t</i> test limits the ability to model time trends and handle missing data</li> <li>No mediation or moderation analyses conducted</li> </ul>	<ul style="list-style-type: none"> <li>Use linear mixed model in the future to model repeated measures [59]</li> <li>Explore mediation pathways (eg, goal → engagement → PA) [60]</li> </ul>

<sup>a</sup>PA: physical activity.

<sup>b</sup>RCT: randomized controlled trial.

<sup>c</sup>Participants selected 'Caucasian' in the original questionnaire; this has been reported as 'White' in the manuscript to reflect current terminology.

Despite fairly high engagement and acceptability, the lack of a significant increase in weekly mean daily step count may reflect a combination of declining adherence over time, relatively high baseline activity levels, and aspects of the intervention design that limited short-term gains in PA. The sample's mean baseline step count (5598, SD 2664 steps/day) was slightly higher than the UK population average (5444 steps/day), suggesting participants may have been more physically active at enrollment than the general population. This raises the possibility of a ceiling effect, whereby participants had less capacity for observable improvement over a short intervention period, which may partly explain the absence of a statistically significant increase in PA. Several observations warrant further consideration. First, no immediate increase in PA was observed from baseline to study week 1, contrary to findings from earlier FI interventions where behavioral effects influence PA behavior immediately and tend to decrease gradually [24,25]. This observed delay may reflect the time required for users to build awareness, navigate app features, and begin integrating new routines [24]. In a future ORBIT phase IIb study, this initial delay could be addressed by incorporating a structured onboarding process that includes guided app walkthroughs, early incentives for app exploration, and prompts to encourage immediate goal engagement. These strategies may help accelerate user activation, reduce the learning curve, and facilitate earlier behavior change. Second, the apparent increase in PA among low-active users and decrease among high-active users should be interpreted cautiously. These subgroup trends are hypothesis-generating and may primarily reflect regression to the mean (ie, a statistical tendency for extreme values to move closer to the average over time) rather than a true intervention effect, particularly in the absence of inferential statistical testing [61,62]. A future ORBIT phase IIb study may protect against this alternative explanation by including a comparison group to better distinguish intervention effects from natural variability. Third, weekly step goal achievement was initially high but declined over the 5-week intervention. Declining goal achievement may have been due to fatigue (ie, physical capability), poorer weather (ie, opportunity), or decreasing perceived value of incentives over time (ie, motivation), and importantly, the mechanics of goal recalibration, which may have unintentionally made goals more difficult over time. A future ORBIT phase IIb trial may address these factors by incorporating adaptive intervention strategies (eg, offering users the option to modify goals during periods of low capability [eg, illness

and fatigue]), integrating weather-responsive prompts or indoor activity suggestions to mitigate environmental barriers, and applying variable incentive structures (eg, loss-framed) to sustain motivation over time. These refinements would help test the intervention under more dynamic, real-world conditions and better assess its potential for long-term adherence.

### Comparison With Prior Work

Findings should also be considered in light of similar literature. The current study's recruitment rate ( $\approx 15\%$  [46/285]) was lower than rates reported in similar feasibility trials of digital health interventions promoting PA with microincentives, which have ranged from 21% to 78% [23,26,63]. Several factors may have contributed to this suboptimal conversion to study participation, including passive recruitment strategies (eg, social media campaigns), limited diversity in reward offerings, and participant burden associated with a multistep consent process. Nonetheless, the observed retention rate (about half) and level of engagement (40% [18/46] of users completing  $\geq 50\%$  of available quizzes [ $\geq 12$  of 24 quizzes]) were higher than rates typically reported with prior digital health interventions, where engagement, more broadly defined as state-like construct that refers to the extent (eg, amount, frequency, duration, and depth) of usage, typically ranges between 20% and 30% [58]. Notably, higher app engagement has been linked to greater intervention effects [8,38], suggesting that the fairly high engagement rate observed here (40%) is encouraging from a health behavior change perspective. Favorable exit survey responses align with previous evidence suggesting that personalized designs (eg, personalized daily step goals) and small, immediate microincentives enhance intervention acceptability [24,64]. Importantly, the sample included a relatively high proportion of low-income participants (18.9%) compared to the national average in the United Kingdom (13%). This is notable given that low-income populations often encounter more barriers to PA engagement compared to high-income groups [65]. Although significant PA changes were not observed in this feasibility study, exploratory subgroup descriptive summaries suggest potential among higher-risk groups (eg, inactive participants and those living with chronic conditions). These possible trends are consistent with prior research indicating that microincentives may be particularly effective in promoting PA among at-risk populations [13,18].

## Limitations

This study has several strengths. Step data was collected using validated step tracking devices (ie, smartphones and fitness trackers in free-living conditions) [66], and the app's cost-free design may have appealed more to lower-income, higher-risk users. Nonetheless, several limitations warrant consideration. First, this study's low recruitment rate limited the sample size and statistical power to detect meaningful PA changes, as well as sample diversity. The requirement for 2 separate consents (ie, one through the app's onboarding system [n=46] and another within the app for step count data collection [n=22]) likely reduced participant recruitment. This 2-stage structure reflected a separation between consent for data collection of standard app use and consent for step count-specific data collection, a design that was partly inherent to the app's real-world deployment. However, future trials may benefit from integrating or streamlining consent procedures to minimize participant burden and improve enrollment efficiency in digital health interventions. One practical approach is the adoption of a tiered, integrated digital consent model, in which participants provide initial consent for minimal, low-risk data collection (eg, nonidentifiable engagement metrics) during onboarding, followed by feature-specific (eg, linking a wearable device), just-in-time consent prompts delivered contextually. This sand-box-style consent approach maintains ethical transparency and participant autonomy while reducing upfront friction and represents an appropriate, measurable methodological refinement suitable for evaluation in a subsequent ORBIT phase IIb study. Second, rolling participant recruitment meant the approximately 6-week evaluation window varied somewhat by participant (eg, some participant data collected in temperate September-October and some in cooler November-December), potentially introducing seasonal (eg, weather) and history bias [64]. Third, if participants did not carry their smartphones or wearable devices consistently throughout the days and weeks, there may be gaps in the data recorded at baseline and during the intervention period [66]. These gaps may lead to an underestimation of the actual number of steps taken per day (ie, measurement bias). However, incentivization may have addressed this by increasing the likelihood that participants consistently carry their devices during the intervention. Fourth, this feasibility study used a single-arm design without a control group, limiting causal inference. Without a comparator, it is difficult to determine whether observed outcomes are attributable to the intervention itself or to external factors (eg, seasonal changes and concurrent health campaigns). Fifth, although the current study's approach to assessing engagement aligns with recommended practices for reporting mHealth usage data (eg, macro-level of engagement) [67] and user-reported

acceptability (eg, exit survey responses) [31], the proxy used may not fully capture the complexity of user engagement [68]. Alternative metrics (eg, frequency, duration, and other feature use) could have provided a more nuanced understanding. Such depth is important for identifying which components were most engaging, informing iterative improvements, enhancing user retention, and elucidating potential mechanisms linking engagement to PA outcomes. Without these insights, key patterns (eg, points of disengagement) may remain undetected [68]. Sixth, although the target behavior chosen for this app was to significantly increase users' daily steps, it is important to note that there are other behaviors related to movement that could be targeted. For example, in addition to encouraging users to walk or run more, the app could also support users to be less sedentary (ie, sit less) or to increase other moderate-to-vigorous intensity activities (eg, cycling to work). Seventh, although step goals were recalculated weekly based on prior step data (eg, median step count from the previous 7-day week), no explicit minimum or maximum thresholds were applied; goals were determined entirely by participant data, without the use of predefined minimum or maximum thresholds or manual adjustments. This may have contributed to potential plateau effects in participant motivation and engagement over time [69]. Eighth, participants exhibited relatively high baseline engagement with PA and self-monitoring behaviors. Mean baseline daily step counts exceeded national averages, and 65.9% of participants reported using a fitness tracker at baseline (ie, nearly double the UK population prevalence [ $\approx 35\%$ ]). This suggests that the sample may have been more motivated, health-conscious, or technologically engaged than the general population. Such preexisting engagement may have positively influenced app uptake, sustained use, and favorable acceptability ratings, and therefore, the high engagement and acceptability observed may not fully generalize to populations with lower baseline activity levels or less familiarity or interest in digital health technologies. Finally, in its current form, Caterpillar is not suitable for people who cannot adequately read and comprehend English, limiting its generalizability, particularly in diverse urban settings such as Leeds, where nearly 10% of adults report a first language other than English [70].

## Conclusions

This ORBIT model phase IIa study served to assess the feasibility of the intervention and refine the delivery and examination of the Caterpillar app that offered microincentives to boost PA in the United Kingdom. Relatively high app engagement and acceptability, and preliminary evidence of PA increases, support advancing to a more robust (ie, 2-arm) pilot evaluation.

## Acknowledgments

The authors would like to express their gratitude to the users of the Caterpillar app for their valuable feedback and participation in the development and improvement of the app. The contributions of the Active Leeds City Council and National Health Service health guidelines are also acknowledged for providing expert health information content throughout the course of this project. Appreciation is further extended to the technical team for their hard work and dedication in bringing the app to

fruition. No generative AI tools (eg, ChatGPT [OpenAI], Gemini [Google LLC], or similar) were used to write or revise this paper. All content, including survey questions, data analysis, and interpretations, was developed solely by this study's team.

### Funding

The study was funded by merit-based scholarship through the Ontario Graduate Scholarship (OGS) Program, provided by the Province of Ontario (Ministry of Colleges and Universities), and administered by Western University for the purposes of conducting research. The funder had no involvement in this study's design, data collection, analysis, interpretation, or the writing of this paper.

### Data Availability

The datasets generated and analyzed during this study are available upon reasonable request.

### Authors' Contributions

BS and MSM conceptualized and designed this study, drafted the initial paper, coordinated and supervised data collection, analyzed the data, and reviewed and revised this paper. MSH, HP, LMV, and MSM critically revised this paper. BS was responsible for the overall content as guarantor. All authors (BS, MSH, HP, LMV, and MSM) approved this final paper as submitted and agree to be accountable for all aspects of this work.

### Conflicts of Interest

MSH received a 50/50 grant from Caterpillar and the Mitacs Accelerate Grant to be able to conduct research and work in an industry capacity. MSM (principal investigator) provided unpaid consulting services to Caterpillar Health Inc between January 2021 and January 2023 (≈10 hours/month). The company agreed to provide funding in the form of a course buyout during a previous academic year (September 2021 to May 2022; CAD \$15,000 [US \$10,924.50]) to free up time for MSM to continue to provide such service. The results of this study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation. This is approved by the Faculty of Health Sciences dean, Jayne Garland, PhD.

### Multimedia Appendix 1

Demographic and health survey, exit survey questions, and responses.

[\[DOCX File \(Microsoft Word File\), 28 KB-Multimedia Appendix 1\]](#)

### Multimedia Appendix 2

Caterpillar health BCTs (behavior change technique) and COM-B (capability, opportunity, motivation-behavior).

[\[DOCX File \(Microsoft Word File\), 39 KB-Multimedia Appendix 2\]](#)

### Multimedia Appendix 3

Baseline characteristics of subgroup samples.

[\[DOCX File \(Microsoft Word File\), 36 KB-Multimedia Appendix 3\]](#)

### Checklist 1

STROBE statement checklist.

[\[PDF File \(Adobe File\), 10288 KB-Checklist 1\]](#)

### References

1. Physical activity: applying all our health. Gov.UK. Mar 10, 2022. URL: <https://www.gov.uk/government/publications/physical-activity-applying-all-our-health> [Accessed 2026-03-21]
2. Gao Z, Lee JE. Emerging technology in promoting physical activity and health: challenges and opportunities. *J Clin Med*. Nov 1, 2019;8(11):1830. [doi: [10.3390/jcm8111830](https://doi.org/10.3390/jcm8111830)] [Medline: [31683951](https://pubmed.ncbi.nlm.nih.gov/31683951/)]
3. Mobile fact sheet. Pew Research Center. Nov 20, 2025. URL: <https://www.pewresearch.org/internet/fact-sheet/mobile/> [Accessed 2026-03-21]
4. Baker N. UK mobile phone statistics. Uswitch. Aug 14, 2025. URL: <https://www.uswitch.com/mobiles/studies/mobile-statistics> [Accessed 2026-03-21]
5. Aldenaini N, Oyebo O, Orji R, Sampalli S. Mobile phone-based persuasive technology for physical activity and sedentary behavior: a systematic review. *Front Comput Sci*. Jul 24, 2020;2(2):673-681. [doi: [10.3389/fcomp.2020.00019](https://doi.org/10.3389/fcomp.2020.00019)]
6. Kamel Boulos MN, Yang SP. Mobile physical activity planning and tracking: a brief overview of current options and desiderata for future solutions. *mHealth*. 2021;7:13. [doi: [10.21037/mhealth.2020.01.01](https://doi.org/10.21037/mhealth.2020.01.01)] [Medline: [33634196](https://pubmed.ncbi.nlm.nih.gov/33634196/)]
7. Mateo GF, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. *J Med Internet Res*. Nov 10, 2015;17(11):e253. [doi: [10.2196/jmir.4836](https://doi.org/10.2196/jmir.4836)] [Medline: [26554314](https://pubmed.ncbi.nlm.nih.gov/26554314/)]

8. Boonmanunt S, Pattanaprteep O, Ongphiphadhanakul B, et al. Evaluation of the effectiveness of behavioral economic incentive programs for goal achievement on healthy diet, weight control and physical activity: a systematic review and network meta-analysis. *Ann Behav Med.* Apr 12, 2023;57(4):277-287. [doi: [10.1093/abm/kaac066](https://doi.org/10.1093/abm/kaac066)] [Medline: [36367428](https://pubmed.ncbi.nlm.nih.gov/36367428/)]
9. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. *J Med Internet Res.* Jun 29, 2017;19(6):e232. [doi: [10.2196/jmir.7126](https://doi.org/10.2196/jmir.7126)] [Medline: [28663162](https://pubmed.ncbi.nlm.nih.gov/28663162/)]
10. Krath J, Schürmann L, von Korfflesch HFO. Revealing the theoretical basis of gamification: a systematic review and analysis of theory in research on gamification, serious games and game-based learning. *Comput Human Behav.* Dec 2021;125:106963. [doi: [10.1016/j.chb.2021.106963](https://doi.org/10.1016/j.chb.2021.106963)]
11. O'Donoghue T, Rabin M. Present bias: lessons learned and to be learned. *Am Econ Rev.* May 1, 2015;105(5):273-279. [doi: [10.1257/aer.p20151085](https://doi.org/10.1257/aer.p20151085)]
12. Eberth B, van der Pol M, Kemenev D. Is time preference and present bias associated with the decision to start, quit or maintain physical activity over time? *J Public Health (Berl).* Mar 2022;30(3):587-597. [doi: [10.1007/s10389-020-01334-8](https://doi.org/10.1007/s10389-020-01334-8)]
13. Luong MLN, Hall M, Bennell KL, Kasza J, Harris A, Hinman RS. The impact of financial incentives on physical activity: a systematic review and meta-analysis. *Am J Health Promot.* Feb 2021;35(2):236-249. [doi: [10.1177/0890117120940133](https://doi.org/10.1177/0890117120940133)] [Medline: [32672050](https://pubmed.ncbi.nlm.nih.gov/32672050/)]
14. Mitchell MS, Orstad SL, Biswas A, et al. Financial incentives for physical activity in adults: systematic review and meta-analysis. *Br J Sports Med.* Nov 2020;54(21):1259-1268. [doi: [10.1136/bjsports-2019-100633](https://doi.org/10.1136/bjsports-2019-100633)] [Medline: [31092399](https://pubmed.ncbi.nlm.nih.gov/31092399/)]
15. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. *Am J Health Promot.* 1997;12(1):38-48. [doi: [10.4278/0890-1171-12.1.38](https://doi.org/10.4278/0890-1171-12.1.38)] [Medline: [10170434](https://pubmed.ncbi.nlm.nih.gov/10170434/)]
16. Rondina II R, Pearson EK, Prapavessis H, White L, Richard S, Mitchell MS. Bright spots, physical activity investments that (almost) worked: Carrot Rewards app, driving engagement with pennies a day. *Br J Sports Med.* Aug 2020;54(15):927-929. [doi: [10.1136/bjsports-2019-100969](https://doi.org/10.1136/bjsports-2019-100969)] [Medline: [31712256](https://pubmed.ncbi.nlm.nih.gov/31712256/)]
17. Department of Health & Social Care. Gov.UK. URL: <https://www.gov.uk/government/organisations/department-of-health-and-social-care> [Accessed 2026-03-21]
18. Rohde KIM, Verbeke W. We like to see you in the gym—a field experiment on financial incentives for short and long term gym attendance. *J Econ Behav Organ.* Feb 2017;134:388-407. [doi: [10.1016/j.jebo.2016.12.012](https://doi.org/10.1016/j.jebo.2016.12.012)]
19. Leahey TM, Subak LL, Fava J, et al. Benefits of adding small financial incentives or optional group meetings to a web-based statewide obesity initiative. *Obesity (Silver Spring).* Jan 2015;23(1):70-76. [doi: [10.1002/oby.20937](https://doi.org/10.1002/oby.20937)] [Medline: [25384463](https://pubmed.ncbi.nlm.nih.gov/25384463/)]
20. Condliffe S, Işgın E, Fitzgerald B. Get thee to the gym! A field experiment on improving exercise habits. *J Behav Exp Econ.* Oct 2017;70:23-32. [doi: [10.1016/j.socec.2017.07.007](https://doi.org/10.1016/j.socec.2017.07.007)]
21. Matsuoka Y, Yoshida H, Hanazato MA. A smartphone-based shopping mall walking program and daily walking steps. *JAMA Netw Open.* Jan 2, 2024;7(1):e2353957. [doi: [10.1001/jamanetworkopen.2023.53957](https://doi.org/10.1001/jamanetworkopen.2023.53957)] [Medline: [38289599](https://pubmed.ncbi.nlm.nih.gov/38289599/)]
22. Kamada M, Hayashi H, Shiba K, et al. Large-scale fandom-based gamification intervention to increase physical activity: a quasi-experimental study. *Med Sci Sports Exerc.* Jan 1, 2022;54(1):181-188. [doi: [10.1249/MSS.0000000000002770](https://doi.org/10.1249/MSS.0000000000002770)] [Medline: [34366420](https://pubmed.ncbi.nlm.nih.gov/34366420/)]
23. Carrera M, Royer H, Stehr M, Sydnor J. Can financial incentives help people trying to establish new habits? Experimental evidence with new gym members. *J Health Econ.* Mar 2018;58:202-214. [doi: [10.1016/j.jhealeco.2018.02.010](https://doi.org/10.1016/j.jhealeco.2018.02.010)] [Medline: [29550665](https://pubmed.ncbi.nlm.nih.gov/29550665/)]
24. Mitchell M, Lau E, White L, Faulkner G. Commercial app use linked with sustained physical activity in two Canadian provinces: a 12-month quasi-experimental study. *Int J Behav Nutr Phys Act.* Feb 25, 2020;17(1):24. [doi: [10.1186/s12966-020-00926-7](https://doi.org/10.1186/s12966-020-00926-7)] [Medline: [32098625](https://pubmed.ncbi.nlm.nih.gov/32098625/)]
25. Pearson E, Prapavessis H, Higgins C, Petrella R, White L, Mitchell M. Adding team-based financial incentives to the Carrot Rewards physical activity app increases daily step count on a population scale: a 24-week matched case control study. *Int J Behav Nutr Phys Act.* Nov 19, 2020;17(1):139. [doi: [10.1186/s12966-020-01043-1](https://doi.org/10.1186/s12966-020-01043-1)] [Medline: [33208166](https://pubmed.ncbi.nlm.nih.gov/33208166/)]
26. Willms A, Rhodes RE, Liu S. Effects of mobile-based financial incentive interventions for adults at risk of developing hypertension: feasibility randomized controlled trial. *JMIR Form Res.* Mar 24, 2023;7:e36562. [doi: [10.2196/36562](https://doi.org/10.2196/36562)] [Medline: [36961486](https://pubmed.ncbi.nlm.nih.gov/36961486/)]
27. Czajkowski SM, Powell LH, Adler N, et al. From ideas to efficacy: the ORBIT model for developing behavioral treatments for chronic diseases. *Health Psychol.* Oct 2015;34(10):971-982. [doi: [10.1037/hea0000161](https://doi.org/10.1037/hea0000161)] [Medline: [25642841](https://pubmed.ncbi.nlm.nih.gov/25642841/)]
28. Health inequality monitor. World Health Organization; Apr 19, 2023. URL: [https://www.who.int/data/inequality-monitor/assessment\\_toolkit](https://www.who.int/data/inequality-monitor/assessment_toolkit) [Accessed 2026-03-21]

29. Healthy weight among adults. Hull City Council. URL: <https://www.hulljsna.com/adults/lifestyle-factors-adults/healthy-weight-adults/> [Accessed 2026-03-31]
30. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med.* Oct 16, 2007;147(8):573-577. [doi: [10.7326/0003-4819-147-8-200710160-00010](https://doi.org/10.7326/0003-4819-147-8-200710160-00010)] [Medline: [17938396](https://pubmed.ncbi.nlm.nih.gov/17938396/)]
31. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR mHealth uHealth.* Mar 11, 2015;3(1):e27. [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
32. Terhorst Y, Philippi P, Sander LB, et al. Validation of the Mobile Application Rating Scale (MARS). *PLoS One.* 2020;15(11):e0241480. [doi: [10.1371/journal.pone.0241480](https://doi.org/10.1371/journal.pone.0241480)] [Medline: [33137123](https://pubmed.ncbi.nlm.nih.gov/33137123/)]
33. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci.* Apr 23, 2011;6(1):42. [doi: [10.1186/1748-5908-6-42](https://doi.org/10.1186/1748-5908-6-42)] [Medline: [21513547](https://pubmed.ncbi.nlm.nih.gov/21513547/)]
34. Michie S, Richardson M, Johnston M, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med.* Aug 2013;46(1):81-95. [doi: [10.1007/s12160-013-9486-6](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]
35. Michie S, Wood CE, Johnston M, et al. Behaviour change techniques: the development and evaluation of a taxonomic method for reporting and describing behaviour change interventions (BCTTv1). *Health Psychol Rev.* 2015;9(1):1-11. [doi: [10.3310/hta19990](https://doi.org/10.3310/hta19990)]
36. Tudor-Locke C, Craig CL, Brown WJ, et al. How many steps/day are enough? For adults. *Int J Behav Nutr Phys Act.* Jul 28, 2011;8:79. [doi: [10.1186/1479-5868-8-79](https://doi.org/10.1186/1479-5868-8-79)] [Medline: [21798015](https://pubmed.ncbi.nlm.nih.gov/21798015/)]
37. Census. Office for National Statistics. URL: <https://www.ons.gov.uk/census> [Accessed 2026-03-21]
38. Althoff T, Sosič R, Hicks JL, King AC, Delp SL, Leskovec J. Large-scale physical activity data reveal worldwide activity inequality. *Nat New Biol.* Jul 20, 2017;547(7663):336-339. [doi: [10.1038/nature23018](https://doi.org/10.1038/nature23018)] [Medline: [28693034](https://pubmed.ncbi.nlm.nih.gov/28693034/)]
39. Adams MA, Todd M, Angadi SS, et al. Adaptive goals and reinforcement timing to increase physical activity in adults: a factorial randomized trial. *Am J Prev Med.* Feb 2022;62(2):e57-e68. [doi: [10.1016/j.amepre.2021.09.014](https://doi.org/10.1016/j.amepre.2021.09.014)] [Medline: [35000693](https://pubmed.ncbi.nlm.nih.gov/35000693/)]
40. Bikou AG, Deligianni E, Dermiki-Gkana F, et al. Improving participant recruitment in clinical trials: comparative analysis of innovative digital platforms. *J Med Internet Res.* Dec 18, 2024;26:e60504. [doi: [10.2196/60504](https://doi.org/10.2196/60504)] [Medline: [39693619](https://pubmed.ncbi.nlm.nih.gov/39693619/)]
41. Klein D, Montgomery A, Begale M, et al. Building a digital health research platform to enable recruitment, enrollment, data collection, and follow-up for a highly diverse longitudinal US cohort of 1 million people in the all of US research program: design and implementation study. *J Med Internet Res.* Jan 15, 2025;27:e60189. [doi: [10.2196/60189](https://doi.org/10.2196/60189)] [Medline: [39813673](https://pubmed.ncbi.nlm.nih.gov/39813673/)]
42. Adams MA, Hurley JC, Phillips CB, et al. Rationale, design, and baseline characteristics of WalkIT Arizona: a factorial randomized trial testing adaptive goals and financial reinforcement to increase walking across higher and lower walkable neighborhoods. *Contemp Clin Trials.* Jun 2019;81:87-101. [doi: [10.1016/j.cct.2019.05.001](https://doi.org/10.1016/j.cct.2019.05.001)] [Medline: [31063868](https://pubmed.ncbi.nlm.nih.gov/31063868/)]
43. Lev V, Oppezzo MA. Measuring intra-individual physical activity variability using consumer-grade activity devices. *Front Digit Health.* 2023;5:1239759. [doi: [10.3389/fdgth.2023.1239759](https://doi.org/10.3389/fdgth.2023.1239759)] [Medline: [37744687](https://pubmed.ncbi.nlm.nih.gov/37744687/)]
44. Fitzpatrick SE, Reddy S, Lommel TS, et al. Physical activity and physical function improved following a community-based intervention in older adults in Georgia senior centers. *J Nutr Elder.* 2008;27(1-2):135-154. [doi: [10.1080/01639360802060223](https://doi.org/10.1080/01639360802060223)] [Medline: [18928194](https://pubmed.ncbi.nlm.nih.gov/18928194/)]
45. Sallis JF, Cervero RB, Ascher W, Henderson KA, Kraft MK, Kerr J. An ecological approach to creating active living communities. *Annu Rev Public Health.* 2006;27:297-322. [doi: [10.1146/annurev.publhealth.27.021405.102100](https://doi.org/10.1146/annurev.publhealth.27.021405.102100)] [Medline: [16533119](https://pubmed.ncbi.nlm.nih.gov/16533119/)]
46. Spencer L, Adams TB, Malone S, Roy L, Yost E. Applying the transtheoretical model to exercise: a systematic and comprehensive review of the literature. *Health Promot Pract.* Oct 2006;7(4):428-443. [doi: [10.1177/1524839905278900](https://doi.org/10.1177/1524839905278900)] [Medline: [16840769](https://pubmed.ncbi.nlm.nih.gov/16840769/)]
47. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health Psychol.* Nov 2009;28(6):690-701. [doi: [10.1037/a0016136](https://doi.org/10.1037/a0016136)] [Medline: [19916637](https://pubmed.ncbi.nlm.nih.gov/19916637/)]
48. Jauho AM, Pyky R, Ahola R, et al. Effect of wrist-worn activity monitor feedback on physical activity behavior: a randomized controlled trial in Finnish young men. *Prev Med Rep.* 2015;2:628-634. [doi: [10.1016/j.pmedr.2015.07.005](https://doi.org/10.1016/j.pmedr.2015.07.005)] [Medline: [26844128](https://pubmed.ncbi.nlm.nih.gov/26844128/)]

49. Barte JCM, Wendel-Vos GCW. A systematic review of financial incentives for physical activity: the effects on physical activity and related outcomes. *Behav Med*. 2017;43(2):79-90. [doi: [10.1080/08964289.2015.1074880](https://doi.org/10.1080/08964289.2015.1074880)] [Medline: [26431076](https://pubmed.ncbi.nlm.nih.gov/26431076/)]
50. Higgins TJ, Middleton KR, Winner L, Janelle CM. Physical activity interventions differentially affect exercise task and barrier self-efficacy: a meta-analysis. *Health Psychol*. Aug 2014;33(8):891-903. [doi: [10.1037/a0033864](https://doi.org/10.1037/a0033864)] [Medline: [23957904](https://pubmed.ncbi.nlm.nih.gov/23957904/)]
51. Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. *J Med Internet Res*. Aug 15, 2014;16(8):e192. [doi: [10.2196/jmir.3469](https://doi.org/10.2196/jmir.3469)] [Medline: [25131661](https://pubmed.ncbi.nlm.nih.gov/25131661/)]
52. Langer SL, Castro FG, Chen ACC, et al. Recruitment and retention of underrepresented and vulnerable populations to research. *Public Health Nurs*. Nov 2021;38(6):1102-1115. [doi: [10.1111/phn.12943](https://doi.org/10.1111/phn.12943)] [Medline: [34240459](https://pubmed.ncbi.nlm.nih.gov/34240459/)]
53. Pocock SJ, Clayton TC, Stone GW. Design of major randomized trials. *J AM Coll Cardiol*. Dec 2015;66(24):2757-2766. [doi: [10.1016/j.jacc.2015.10.036](https://doi.org/10.1016/j.jacc.2015.10.036)] [Medline: [26700838](https://pubmed.ncbi.nlm.nih.gov/26700838/)]
54. Alkhalidi G, Modrow K, Hamilton F, Pal K, Ross J, Murray E. Promoting engagement with a digital health intervention (HeLP-Diabetes) using email and text message prompts: mixed-methods study. *Interact J Med Res*. Aug 22, 2017;6(2):e14. [doi: [10.2196/ijmr.6952](https://doi.org/10.2196/ijmr.6952)] [Medline: [28829328](https://pubmed.ncbi.nlm.nih.gov/28829328/)]
55. Bender JL, Akinnibosun R, Puri N, et al. A comparison of the sociodemographic, medical, and psychosocial characteristics of adolescents and young adults diagnosed with cancer recruited in-person and online: a Canadian cross-sectional survey. *Digit Health*. 2023;9:20552076231205278. [doi: [10.1177/20552076231205278](https://doi.org/10.1177/20552076231205278)] [Medline: [37900258](https://pubmed.ncbi.nlm.nih.gov/37900258/)]
56. Aguilar-Farias N, Martino-Fuentealba P, Salom-Diaz N, Brown WJ. How many days are enough for measuring weekly activity behaviours with the ActivPAL in adults? *J Sci Med Sport*. Jun 2019;22(6):684-688. [doi: [10.1016/j.jsams.2018.12.004](https://doi.org/10.1016/j.jsams.2018.12.004)] [Medline: [30579814](https://pubmed.ncbi.nlm.nih.gov/30579814/)]
57. Fjeldsoe B, Neuhaus M, Winkler E, Eakin E. Systematic review of maintenance of behavior change following physical activity and dietary interventions. *Health Psychol*. Jan 2011;30(1):99-109. [doi: [10.1037/a0021974](https://doi.org/10.1037/a0021974)] [Medline: [21299298](https://pubmed.ncbi.nlm.nih.gov/21299298/)]
58. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. *Transl Behav Med*. Jun 2017;7(2):254-267. [doi: [10.1007/s13142-016-0453-1](https://doi.org/10.1007/s13142-016-0453-1)] [Medline: [27966189](https://pubmed.ncbi.nlm.nih.gov/27966189/)]
59. West BT, Welch KB, Galecki AT. *Linear Mixed Models: A Practical Guide Using Statistical Software*. 2nd ed. CRC Press; 2014. [doi: [10.1201/b17198](https://doi.org/10.1201/b17198)]
60. Hayes AF. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. 2nd ed. Guilford Press; 2018. ISBN: 9781462549030
61. Beets MW, von Klingraeff L, Weaver RG, Armstrong B, Burkart S. Small studies, big decisions: the role of pilot/feasibility studies in incremental science and premature scale-up of behavioral interventions. *Pilot Feasibility Stud*. Sep 10, 2021;7(1):173. [doi: [10.1186/s40814-021-00909-w](https://doi.org/10.1186/s40814-021-00909-w)] [Medline: [34507624](https://pubmed.ncbi.nlm.nih.gov/34507624/)]
62. König LM, Pasko K, Baga K, Harsora R, Arigo D. Isolating the role of researcher observation on reactivity to the measurement of physical activity. *Appl Psychol Health Well Being*. Feb 2025;17(1):e12630. [doi: [10.1111/aphw.12630](https://doi.org/10.1111/aphw.12630)] [Medline: [39703095](https://pubmed.ncbi.nlm.nih.gov/39703095/)]
63. Klasnja P, Smith S, Seewald NJ, et al. Efficacy of contextually tailored suggestions for physical activity: a micro-randomized optimization trial of HeartSteps. *Ann Behav Med*. May 3, 2019;53(6):573-582. [doi: [10.1093/abm/kay067](https://doi.org/10.1093/abm/kay067)] [Medline: [30192907](https://pubmed.ncbi.nlm.nih.gov/30192907/)]
64. Derlyatka A, Fomenko O, Eck F, Khmelev E, Elliott MT. Bright spots, physical activity investments that work: Sweatcoin: a steps generated virtual currency for sustained physical activity behaviour change. *Br J Sports Med*. Sep 2019;53(18):1195-1196. [doi: [10.1136/bjsports-2018-099739](https://doi.org/10.1136/bjsports-2018-099739)] [Medline: [30609999](https://pubmed.ncbi.nlm.nih.gov/30609999/)]
65. Mendoza-Vasquez AS, Linke S, Muñoz M, et al. Promoting physical activity among underserved populations. *Curr Sports Med Rep*. 2016;15(4):290-297. [doi: [10.1249/JSR.0000000000000276](https://doi.org/10.1249/JSR.0000000000000276)] [Medline: [27399827](https://pubmed.ncbi.nlm.nih.gov/27399827/)]
66. Yao J, Tan CS, Lim N, Tan J, Chen C, Müller-Riemenschneider F. Number of daily measurements needed to estimate habitual step count levels using wrist-worn trackers and smartphones in 212,048 adults. *Sci Rep*. May 5, 2021;11(1):9633. [doi: [10.1038/s41598-021-89141-3](https://doi.org/10.1038/s41598-021-89141-3)] [Medline: [33953288](https://pubmed.ncbi.nlm.nih.gov/33953288/)]
67. Short CE, DeSmet A, Woods C, et al. Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies. *J Med Internet Res*. Nov 16, 2018;20(11):e292. [doi: [10.2196/jmir.9397](https://doi.org/10.2196/jmir.9397)] [Medline: [30446482](https://pubmed.ncbi.nlm.nih.gov/30446482/)]
68. Perski O, Blandford A, Garnett C, Crane D, West R, Michie S. A self-report measure of engagement with digital behavior change interventions (DBCIs): development and psychometric evaluation of the “DBCI Engagement Scale”. *Transl Behav Med*. Feb 3, 2020;10(1):267-277. [doi: [10.1093/tbm/ibz039](https://doi.org/10.1093/tbm/ibz039)] [Medline: [30927357](https://pubmed.ncbi.nlm.nih.gov/30927357/)]

69. Epton T, Currie S, Armitage CJ. Unique effects of setting goals on behavior change: systematic review and meta-analysis. *J Consult Clin Psychol*. Dec 2017;85(12):1182-1198. [doi: [10.1037/ccp0000260](https://doi.org/10.1037/ccp0000260)] [Medline: [29189034](https://pubmed.ncbi.nlm.nih.gov/29189034/)]
70. Leeds: main language (detailed). Census data UK. URL: <https://censusdata.uk/e08000035-leeds/ts024-main-language-detailed> [Accessed 2026-03-21]

---

**Abbreviations:**

**BCW:** behavior change wheel

**COM-B:** capability, opportunity, motivation–behavior

**FI:** financial incentive

**KPMG:** Klynveld Peat Marwick Goerdeler

**MD:** mean difference

**mHealth:** mobile health

**ORBIT:** Obesity-Related Behavioral Intervention Trial

**PA:** physical activity

**RCT:** randomized controlled trial

**STROBE:** Strengthening the Reporting of Observational Studies in Epidemiology

---

*Edited by Matthew Balcarras; peer-reviewed by Jinyu Guo, Taofeek Yusuff; submitted 22.Oct.2025; final revised version received 04.Mar.2026; accepted 04.Mar.2026; published 10.Apr.2026*

*Please cite as:*

Salmani B, Hiemstra MS, Prapavessis H, Vanderloo LM, Mitchell MS

*Evaluating an Incentive-Based mHealth App for Physical Activity Promotion Using the Obesity-Related Behavioral Intervention Trial Model: Small Cohort Study*

*JMIR Form Res 2026;10:e85484*

URL: <https://formative.jmir.org/2026/1/e85484>

doi: [10.2196/85484](https://doi.org/10.2196/85484)

© Babac Salmani, Madison S Hiemstra, Harry Prapavessis, Leigh M Vanderloo, Marc S Mitchell. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 10.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.