ann. behav. med. (2021) XX:1–7
DOI: 10.1093/abm/kaab028

BRIEF REPORT

Daily Motivational Text Messages to Promote Physical Activity in University Students: Results From a Microrandomized Trial

Caroline A. Figueroa, MD, PHD[1] · Nina Deliu, MSc[2] · Bibhas Chakraborty, PhD[3,4,5] · Arghavan Modiri, Msc[6] · Jing Xu, PhD[3,7] · Jai Aggarwal, MSc[6] · Joseph Jay Williams, PhD[6] · Courtney Lyles, PhD[8] · Adrian Aguilera, PhD[1,8]

Published online: 19 April 2021

Abstract

Background Low physical activity is an important risk factor for common physical and mental disorders. Physical activity interventions delivered via smartphones can help users maintain and increase physical activity, but outcomes have been mixed.

Purpose Here we assessed the effects of sending daily motivational and feedback text messages in a microrandomized clinical trial on changes in physical activity from one day to the next in a student population.

Methods We included 93 participants who used a physical activity app, "DIAMANTE" for a period of 6 weeks. Every day, their phone pedometer passively tracked participants' steps. They were microrandomized to receive different types of motivational messages, based on a cognitive-behavioral framework, and feedback on their steps. We used generalized estimation equation models to test the effectiveness of feedback and motivational messages on changes in steps from one day to the next.

Results Sending any versus no text message initially resulted in an increase in daily steps (729 steps, $p = .012$), but this effect decreased over time. A multivariate analysis evaluating each text message category separately showed that the initial positive effect was driven by the motivational messages though the effect was small and trend-wise significant (717 steps; $p = .083$), but not the feedback messages (−276 steps, $p = .4$).

Conclusion Sending motivational physical activity text messages based on a cognitive-behavioral framework may have a positive effect on increasing steps, but this decreases with time. Further work is needed to examine using personalization and contextualization to improve the efficacy of text-messaging interventions on physical activity outcomes.

ClinicalTrials.gov Identifier: NCT04440553.

Keywords Mobile health · Microrandomization · Text messaging · Physical activity · Students

✉ Caroline A. Figueroa
c.a.figueroa@berkeley.edu

1 School of Social Welfare, University of California, Berkeley, CA, USA

2 Department of Statistical Sciences, Sapienza University of Rome, Rome, Italy

3 Centre for Quantitative Medicine and Program in Health Services and Systems Research, Duke-NUS Medical School, Singapore, Singapore

4 Department of Statistics and Applied Probability, National University of Singapore, Singapore, Singapore

5 Department of Biostatistics and Bioinformatics, Duke University, Durham, NC, USA

6 Department of Computer Science, University of Toronto, Toronto, ON, Canada

7 Data Science Program, Division of Science and Technology, Beijing Normal University and Hong Kong Baptist University–United International College, Zhuhai, Guangdong, China

8 UCSF Center for Vulnerable Populations, Zuckerberg San Francisco General Hospital, San Francisco, CA, USA

Introduction

Insufficient physical activity (PA) is one of the leading global risk factors of death [1]. It is associated with many common chronic diseases [2] and mental disorders including depression [3]. The World Health Organization recommends 2.5 hr of moderate intensity PA weekly [1]. However, in 2018, less than half of American adults met this goal [4]. Adolescents and university students show even lower PA [5, 6]. There is a need for interventions that help people to increase and maintain PA. Behavioral interventions via mobile devices, such as text messaging and/or smartphone apps, hold great promise for PA promotion. They can help identify the benefits of, and opportunities for, exercise, and aid with goal setting and accountability [7]. Mobile health (mHealth) interventions can also be widely disseminated at relatively low cost [8]. mHealth interventions increase PA, with effect sizes up to 3.10 (Cohen's *d*) after 3-month follow-up (though null effects were also reported) [7, 9–11]. In a systematic review, 16 of 27 reviewed interventions increased PA in university students [12]. In another review, PA interventions decreased depression and anxiety in young people [13].

However, the effects of mHealth PA interventions are mixed [14], and not sustained over longer periods of time [11], which is partly because they do not adapt their messaging content and frequency to participants' changing behaviors [15]. Furthermore, because most mHealth studies evaluate the whole intervention, it is unclear which intervention components most effectively increase daily PA. The microrandomized trial (MRT) is a state-of-the-art experimental design for testing the effects of mHealth the intervention components [16]. In an MRT, individuals are repeatedly randomized to intervention options. This allows researchers to test the effects of separate components, such as categories of text messages (in the current study), on steps within a short time (24 hr in this study).

In an MRT, we tested a mHealth application that sends daily text messages, within categories based on a cognitive-behavioral change model: Capability, Opportunity, Motivation, Behavior (COM-B) [17]. COM-B proposes engaging in a particular behavior depending on the capability (physical and psychological), opportunity (social and physical), and motivation to engage in the behavior more than in other behaviors. Interventions should target at least one of these components for long-lasting behavior change. Researchers have used COM-B to identify PA barriers [18] and design PA interventions [19, 20].

We originally planned to randomize participants to random messaging or a Thompson sampling (TS) condition. TS is a multiarmed bandit technique, used to more intelligently select and personalize messages based on the expected outcome of interest (here increase in steps) [21]. Because of technical difficulties (execution errors in the algorithm), we only enrolled a subset of participants after we fixed the errors in the TS condition. Therefore, we altered the study's main aims by focusing on the MRT. The main aims of this study were (i) to examine the overall effect of sending a text message versus no message over time on PA and (ii) to assess the effectiveness of different types of text messages (motivational and feedback) on PA from one day to the next. Our secondary aims are to examine (i) differences in increases in step count over the course of the study between the TS and uniform random group and (ii) the changes in the depression, anxiety, and behavioral activation (tracks behaviors underlying depression [22]) scores from baseline to follow-up and between groups.

Methods

We use a mHealth app, "DIAMANTE," and a text-messaging platform, HealthySMS, developed by Audacious Software and the authors (https://diamante.healthysms.org/) [23]. DIAMANTE tracks steps pooling from Google Fit, Apple HealthKit, or the built-in phone pedometer, provided that the application remains open. HealthySMS sends text messages and manages participant responses. The clinical trial was prospectively registered on ClinicalTrials.gov Identifier: NCT04440553.

Motivational Text Messages

We designed text messages to fit into the three dimensions of the COM-B model, with a social (i.e., exercising with friends) or an individual connotation (exercising for yourself). We originally designed the messages for a clinical population, but we adapted them by adding messages about the benefits of walking on the brain and concentration and removed messages about chronic disease or family. Participants also received a feedback message on their step count and step goal in the previous day, 2 min apart from the motivational messages.

Experimental Factors

This study is a full factorial design with three factors: Motivational (M, four levels), Feedback Messages (F, five levels), and the Time Frame (T, four levels) (see Fig. 1). Participants received a different combination of M, F, and T daily. Most participants were in a uniform random combination group (microrandomized messages). Participants were unaware of their group membership until the study ended. Supplementary Material shows more information about the TS condition.

Participants

We recruited participants through the Social and Experimental Research Lab (Xlab) app from the University of California, Berkeley, advertised during campus events and on Facebook. Students who did not have a smartphone were unable to exercise due to disability or planned to leave the country during the study were ineligible. The UC Berkeley Committee for Protection of Human Subjects approved the study (ID: 2019-04-12118).

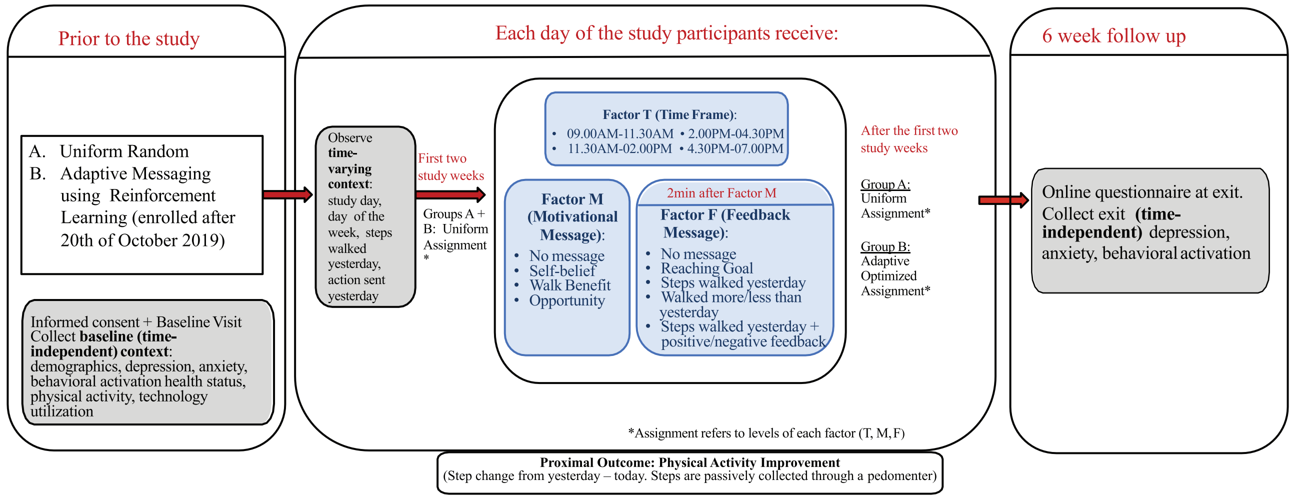

Fig. 1. Study Design.

Study Visits

Participants came into Xlab for informed consent and a baseline survey (outlined below). They received assistance if needed in downloading the app and were instructed to leave the app open. They received $15 USD for the baseline visit and $25 USD for a remote 6-week follow-up survey.

Measures

At baseline and follow-up, students completed a survey on demographics, socioeconomic status, health status, PA, and psychological questionnaires: the Patient Health Questionnaire-8 item (PHQ-8, higher scores indicate greater depression over the past 2 weeks [24]. The PHQ-8 omits the PHQ-9's suicidality question and is preferable to the PHQ-9 in (online) research settings [25]; the General Anxiety Disorder-7 item (GAD-7 [26], higher scores indicate greater anxiety); the Behavioral Activation for Depression–Short Form (BAD-SF, higher scores indicate lower depression risk); and the International PA Questionnaire (IPAQ)−Short Form assesses self-reported PA in the last 6 months [27].

Statistical Analysis

We excluded participants with ≤2 days of data. We conducted a complete case analysis and sensitivity analyses using multiple imputation [28] (see Supplementary Material). We used the generalized estimating equations (GEE) models [29], an extension of generalized linear models and quasi-likelihood estimation methods, widely used in mHealth [30]. We used "geepack" in R [31] and employed an independent working correlation structure (within clusters), taking into account the quasi-information criterion method [32]. For time-dependent covariates, the GEE estimator is consistent under the independent working correlation structure and thus considered a "safe" choice [33].

Main Outcomes

We computed daily step change: today's–yesterday's step count between 0:00 and 23:59. We examined the effect of (i) sending any versus no message on step change; (ii) sending a feedback or motivational message, including the interaction between feedback and motivation; and (iii) the different categories of feedback (k = 4) and motivational (k = 5) messages. All models are adjusted for time (study day). We included the uniform random (n = 66) and TS groups (n = 27, enrolled after October 20, 2019) to increase our sample, but ran sensitivity analyses without the TS group (Supplementary Material).

Secondary Outcomes

We assessed differences in overall PA change between the TS and the uniform random groups. We examined the group effects on changes between baseline and follow-up for PHQ-8 and GAD-7 scores, using a T-test on the change scores because pre/post scores were not normally distributed, and BAD-SF by a two-way repeated-measures ANOVA. Because we did not randomize participants, we did not assess the influence of sociodemographic factors on the adaptive intervention, as originally planned.

Results

One hundred and three students enrolled from September 12, 2019 until October 25, 2019. Seven did not receive the text messages due to technical issues (iOS updates, n = 5, or wrong language setting in the Google

Play store, $n = 2$). Three received messages, but did not transmit data to our server (see flowchart in Fig. 2).

Missing Data

We removed 670 days with missing steps (16%) and excluded three participants with ≤2 days of step data, leaving 93 participants (see Fig. 1 for flowchart and Table 1 for baseline characteristics).

Message Randomization

On average, in 45 days, subjects received a motivational and a feedback message on 27 and 30 days, respectively. Messages were not sent for 17% of the time (700 days) because of system errors. We coded these nonrandomized days as no message (M0 and/or F0). However, we conducted a sensitivity analysis (see Supplementary Material) removing nonrandomized days.

Effects of Text Messages on PA Sending Any Message Versus No Message

When any message was sent (e.g., feedback, motivation, or both), step change increased by 729 ($p = .011$, standardized effect size [δ] 0.147). This effect diminished linearly over time, step change −33 for each day ($p = .004$, $\delta = -0.007$, Table 2).

Motivational and Feedback Messages

A motivational message trend-wise increased steps after correcting for time and message category interactions (717 steps; $p = .083$, $\delta = .144$, Table 3).

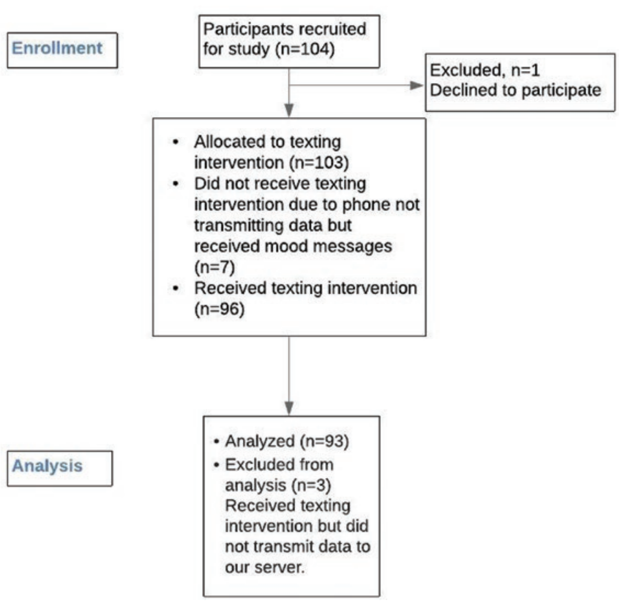

Fig. 2. Flowchart.

Motivational and Feedback Categories (Exploratory)

Motivational messages self-efficacy (414 steps; $\delta = 0.083$, $p = .077$) and opportunity (410 steps; $\delta = 0.083$, $p = .089$) showed trend increases. Steps significantly decreased for feedback on the steps yesterday (−665 steps; $\delta = -0.134$, $p = .002$, Table 4).

Sensitivity Analyses

Supplementary Material shows (post hoc) power analyses and the results above repeated: (i) without the TS group ($n = 27$); (ii) removing all nonrandomized days, and (iii)

Table 1. Baseline characteristics: mean (*SD*) for continuous and *n* (%) for categorical variables

Baseline characteristics	$n = 93$
Female	65 (69.9%)
Male	27 (29%)
Other	1 (1.2%)
Age, mean (*SD*)	20.2 (2.47)
Ethnicity	
Asian or Pacific Islander	51 (54.8%)
Hispanic/Latino(a)	11 (11.8%)
Multiethnic	10 (10.8%)
White or Caucasian	19 (20.4%)
Refused to provide ethnicity	2 (2.4%)
Born in the United States	55 (59.1%)
Engaging in regular physical activity last 6 months[a]	48 (51.6%)
Wants to be more physically active	88 (94.6%)
Self-reported min of moderate/vigorous exercise/week, median, IQR[a]	150 (90/150/171)
PHQ-8 (depressive symptoms, α = 0.93 this sample), mean (*SD*)	5.61 (3.62)
GAD-7 (general anxiety, α this sample 0.77), mean (*SD*)	4.73 (4.84)
BADS-SF (behavioral activation, α = 0.80 this sample, mean (*SD*)	31.1 (8.33)

BAD-SF Behavioral Activation for Depression Scale; *GAD-7* General Anxiety Depression Scale-7; *PHQ-8* Patient Health Questionnaire-8.

[a]Measured by the International Physical Activity Questionnaire.

Table 2. Results of the GEE model, effects of sending any versus no message on steps change

Covariate	Estimate	95% CI	*p*-value
(Intercept)	−685	−1172, −198	.006
Message	729	163, 1295	.012
Study day	27.4	8.53, 46.4	.005
Message × study day	−33.2	−56, −10.4	.004

CI confidence interval; GEE, generalized estimating equations.

Table 3. Results of the GEE model, effects of feedback and motivational messages on steps change

Covariate	Estimate	95% CI	*p*-value
Motivation	717	−93.6, 1527	.083
Feedback	−297	−1089, 496	.463
Day of study	11.9	−3.97, 27.9	.141
Motivation × day of study	−14.6	−44.7, 15.6	.344
Feedback × day of study	−1.65	−31.2, 27.9	.382
Motivation × feedback	−24.4	−676, 627	.570

CI confidence interval; GEE, generalized estimating equations.

Table 4. Results of the GEE models, effects of feedback and motivation categories on step change

Term	Estimate	95% CI	*p*-value
Motivation−Capability	253	−247, 753	.320
Motivation−Self-efficacy	414	−45.4, 874	.077
Motivation−Opportunity	410	−63.2, 883	.089
Feedback−Reaching goal	−147	−703, 408	.603
Feedback−Steps yesterday	−406	−907, 95.5	.113
Feedback−More/less steps than yesterday	−126	−579, 327	.585
Feedback−Steps yesterday + positive/negative	−665	−1082, −247	.002
Study day	1.64	−2.58, 5.85	.447

CI confidence interval; *GEE* generalized estimating equations.

using three imputed datasets. In summary, directions of effects remained similar, but some effects lost significance.

Differences Between TS (n = 27) and Uniform Random Groups (n = 66)

Group membership did not have a significant effect on steps (estimate: −515, confidence interval: −1,536 to 506, p = .32).

Psychological Questionnaires

Eighty-two of 93 subjects provided follow-up data. The PHQ-8 (mean change = 2.68, SD = 3.38, $p < .0001$) and GAD-7 (mean change = 1.18, SD = 3.94, p = .008), but not behavioral activation scores (p = .12), significantly increased from baseline to follow-up. The TS group had a lower change in PHQ-8 scores (p = .0151) than uniform random group (see Supplementary Material).

Discussion

This study examined the effectiveness of motivational text messages on daily step changes in university students. Receiving a message was initially associated with step increases, but this disappeared over time and effects were small (Cohen's $d < 0.2$), Messages may not have tailored contextually (i.e., to participants' daily contexts) and personally (i.e., to fit with a person's personality) enough. Messages that adapt to factors such as time, day of week, and work schedules may be more effective than generic messages [34]. For instance, in an MRT, contextually tailored walking suggestions led to an average increase of 496 daily steps [35]. Comparable to our results however, the effects decreased over time. We had low drop out (3%), but participants may have paid less attention to the messages as the study continued. Future work should focus on measuring and sustaining engagement with (unsupported) texting interventions.

We found that motivational messages may increase PA more than feedback messages. Self-efficacy: the belief that one is capable of behavior change, and opportunity: identifying possibilities for exercise in the current context, showed positive trend effects (exploratory analysis). These two behavior change categories based on the COM-B may be more important than messages about the benefits of exercise for university students—an educated sample who may already be aware of the advantages of exercise. Of note, these effects were small, decreased over time, and disappeared when we removed days that participants were not randomized. This may be due to insufficient power to detect small effects (see our Supplementary Material—post hoc power analyses). Thus, our results need to be confirmed by future work.

Feedback on individuals' steps and goal completion had no effect on step change. Systematic reviews showed that feedback should be actionable (e.g., when, how, and where can you exercise [36]). In addition to feedback, future interventions should therefore give concrete actions for goal completion.

We did not observe differences between the participants who received microrandomized messages, or messaging chosen by a TS algorithm. Our study period may have been too short for the algorithm to start learning, especially with sparse data [21, 37]. To date, mHealth studies using machine learning methods have shown promise, but the small number of studies impedes a rigorous evaluation [38]. Our MRT design examines intervention components (here types of messages), as opposed to the whole intervention. We therefore need RCTs with a long follow-up to assess TS's benefits. To examine this, we are currently conducting a larger RCT in a patient population [23].

The nature of the college semester may explain the unexpected increased depression and anxiety scores from baseline to follow-up. At the end of the study, participants' stress and depression may have increased with exams, which may also have contributed to the messages losing effect over time. Our findings underline that university students are vulnerable for mental health issues [39] and that they may need more tailored mental health support to cope with the pressures of college life during exam periods.

Limitations

This is a convenience sample (university students), with various levels of baseline PA. Although the majority of the participants wanted to exercise more, they may have been more interested in other PA (e.g., the gym or team sports). Furthermore, we experienced technical issues such as our server not sending messages, and missing steps when participants lacked Internet connectivity, or forgot their phones [40]. Furthermore, (small) measurement error related to phone pedometers may influence our findings. We used in-built phone pedometers because of their wide availability—anyone with a smartphone can use our app. They also seem to be as, or more, accurate than wrist or hip worn accelerometers [41]. For instance, the mean absolute percentage of errors was small for iPhone SE and Samsung Galaxy for walking in natural conditions (<3%) compared with a wrist-worn ActiGraph (17%–47%) [41]. However, phone pedometers may underestimate steps [41] and are less accurate when placed further from the body [42]. In addition, as we examine steps over 24 hr, carry-over effects of previous messages may influence step changes. We lack reliable data to assess hourly steps. Because we randomized messaging times, however, we expect that we provide a reasonable approximation of messages' effects on steps.

Conclusion

Sending a motivational text message may be a positive addition to a mHealth PA intervention. However, the effect of the motivational messages was small and diminished over time. Future research should increase personalization, including adapting to participants' daily changing contexts, providing actionable feedback, and taking personal preferences into account.

Supplementary Material

Supplementary material is available at *Annals of Behavioral Medicine* online.

Acknowledgments We thank Chris Karr who developed the DIAMANTE app.

Funding The Experimental Social Science Laboratory (Xlab) at University of California Berkeley funded this study. The parent DIAMANTE RCT is funded by an R01 to Dr. Adrian Aguilera and Dr. Lyles, 1R01 HS25429-01 from the Agency for Healthcare Research and Quality (AHRQ). The sponsors had no role in the study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication. Dr. Chakraborty received support from Duke-NUS Medical School, Singapore, and the Ministry of Education, Singapore.

Compliance With Ethical Standards

Authors' Statement of Conflict of Interest and Adherence to Ethical Standards The authors report no competing interests.

Author Contributions Dr. Figueroa drafted the first version of the manuscript. Ms. Deliu and Dr. Figueroa conducted the statistical analysis. Dr. Chakraborty advised the statistical analysis. Dr. Xu calculated effect sizes and power analyses. Drs. Figueroa, Aguilera, Lyles, and Jay-Williams were involved in the design of the study. Ms. Modiri and Mr. Aggarwal were involved in the design of microrandomization and the Thompson sampling algorithm. All authors contributed to the writing of the final manuscript.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

1. WHO. Physical Activity Fact Sheet. Available at https://www.who.int/news-room/fact-sheets/detail/physical-activity. Accessibility verified May 14, 2019.
2. Brugnara L, Murillo S, Novials A, et al. Low physical activity and its association with diabetes and other cardiovascular risk factors: A nationwide, population-based study. *PLoS One.* 2016;11:e0160959.
3. Choi KW, Chen C-Y, Stein MB, et al. Assessment of bidirectional relationships between physical activity and depression among adults: A 2-Sample Mendelian Randomization Study. *JAMA Psychiatry*. 2019;76(4):399–408.
4. Center for Disease Control and Prevention. Prevalence of self-reported physical inactivity among US Adults by race/ethnicity, state and territory, BRFSS, 2015–2018. 2020. Available at https://www.cdc.gov/physicalactivity/data/inactivity-prevalence-maps/index.html. Accessibility verified February 2020.
5. Kriemler S, Meyer U, Martin E, van Sluijs EM, Andersen LB, Martin BW. Effect of school-based interventions on physical activity and fitness in children and adolescents: A review of reviews and systematic update. *Br J Sports Med.* 2011;45:923–930.
6. Castro O, Bennie J, Vergeer I, Bosselut G, Biddle SJH. How Sedentary are university students? A systematic review and meta-analysis. *Prev Sci.* 2020;21:332–343.
7. Rose T, Barker M, Maria Jacob C, et al. A systematic review of digital interventions for improving the diet and physical activity behaviors of adolescents. *J Adolesc Health.* 2017;61:669–677.
8. Schueller SM, Hunter JF, Figueroa C, Aguilera A. Use of digital mental health for marginalized and underserved populations. *Curr Treat Options Psych*. 2019;6(3):243–255.

9. Roberts AL, Fisher A, Smith L, Heinrich M, Potts HWW. Digital health behaviour change interventions targeting physical activity and diet in cancer survivors: A systematic review and meta-analysis. *J Cancer Surviv.* 2017;11:704–719.
10. Rathbone AL, Prescott J. The use of mobile apps and SMS messaging as physical and mental health interventions: Systematic review. *J Med Internet Res.* 2017;19:e295.
11. Romeo A, Edney S, Plotnikoff R, et al. Can smartphone apps increase physical activity? Systematic review and meta-analysis. *J Med Internet Res.* 2019;21:e12053.
12. Maselli M, Ward PB, Gobbi E, Carraro A. Promoting physical activity among university students: A systematic review of controlled trials. *Am J Health Promot.* 2018;32:1602–1612.
13. Pascoe M, Bailey AP, Craike M, et al. Physical activity and exercise in youth mental health promotion: A scoping review. *BMJ Open Sport Exerc Med*. 2020;6(1).
14. Stuckey MI, Carter SW, Knight E. The role of smartphones in encouraging physical activity in adults. *Int J Gen Med.* 2017;10:293–303.
15. Triantafyllidis A, Kondylakis H, Votis K, Tzovaras D, Maglaveras N, Rahimi K. Features, outcomes, and challenges in mobile health interventions for patients living with chronic diseases: A review of systematic reviews. *Int J Med Inform.* 2019;132:103984.
16. Klasnja P, Hekler EB, Shiffman S, et al. Microrandomized trials: An experimental design for developing just-in-time adaptive interventions. *Health Psychol.* 2015;34S:1220–1228.
17. Michie S, van Stralen MM, West R. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implement Sci.* 2011;6:42.
18. Flannery C, McHugh S, Anaba AE, et al. Enablers and barriers to physical activity in overweight and obese pregnant women: An analysis informed by the theoretical domains framework and COM-B model. *BMC Pregnancy Childbirth.* 2018;18:178.
19. Carney R, Bradshaw T, Yung AR. Physical health promotion for young people at ultra-high risk for psychosis: An application of the COM-B model and behaviour-change wheel. *Int J Ment Health Nurs.* 2016;25:536–545.
20. Nyenhuis S, Ma J, Sharp L. Applying the COM-B model to designing a tailored physical activity intervention for sedentary African American women with asthma. In: B38. Asthma: A panoramic view. *Am J Resp and Crit Care Med.* 2017;195:A3336.
21. Tewari A, Murphy SA. From ads to interventions: Contextual bandits in mobile health. In: Rehg J, Murphy S, Kumar S, eds. *Mobile Health*. Cham, Switzerland: Springer; 2017:495–517.
22. Santos MM. *Validation of the Behavioral Activation for Depression Scale-Short Form (BADS-SF) with Spanish-speaking Latinos*. Theses and Dissertations 155. 2013. Available at https://dc.uwm.edu/etd/155.
23. Aguilera A, Figueroa CA, Hernandez-Ramos R, et al. mHealth app using machine learning to increase physical activity in diabetes and depression: Clinical trial protocol for the DIAMANTE Study. *BMJ Open.* 2020;10:e034723.
24. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: Validity of a brief depression severity measure. *J Gen Intern Med.* 2001;16:606–613.
25. Spangenberg L, Braehler E, Glaesmer H. Identifying depression in the general population-a comparison of PHQ-9, PHQ-8 and PHQ-2. *Zeit Psych Med Psych*. 2012;58(1):3–10.
26. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: The GAD-7. *Arch Intern Med.* 2006;166:1092–1097.
27. Lee PH, Macfarlane DJ, Lam TH, Stewart SM. Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): A systematic review. *Int J Behav Nutr Phys Act.* 2011;8:115.
28. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: What is it and how does it work? *Int J Methods Psychiatr Res.* 2011;20:40–49.
29. Liang K-Y, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika*. 1986;73(1):13–22.
30. Bolger N, Laurenceau J-P. *Intensive Longitudinal Methods: An Introduction to Diary and Experience Sampling Research*. New York, NY: Guilford Press; 2013.
31. Halekoh U, Højsgaard S, Yan J. The R package geepack for generalized estimating equations. *J Stat Softw*. 2006;15(2):1–11.
32. Pan W. Akaike's information criterion in generalized estimating equations. *Biometrics.* 2001;57:120–125.
33. Sullivan Pepe M, Anderson GL. A cautionary note on inference for marginal regression models with longitudinal data and general correlated response data. *Comm Stat-Sim Comp*. 1994;23(4):939–951.
34. Lee MK, Kim J, Forlizzi J, Kiesler S. Personalization revisited: A reflective approach helps people better personalize health services and motivates them to increase physical activity. *Proc ACM Int Conf Ubiquitous Comput*. 2015.
35. 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.* 2019;53:573–582.
36. Schembre SM, Liao Y, Robertson MC, et al. Just-in-time feedback in diet and physical activity interventions: Systematic review and practical design framework. *J Med Internet Res.* 2018;20:e106.
37. Liao P, Greenewald K, Klasnja P, Murphy S. Personalized heartsteps: A reinforcement learning algorithm for optimizing physical activity. *Proc ACM Interact Mob Wearable Ubiquitous Technol*. 2020;4(1):1–22.
38. Triantafyllidis AK, Tsanas A. Applications of machine learning in real-life digital health interventions: Review of the literature. *J Med Internet Res.* 2019;21:e12286.
39. Oswalt SB, Lederer AM, Chestnut-Steich K, Day C, Halbritter A, Ortiz D. Trends in college students' mental health diagnoses and utilization of services, 2009–2015. *J Am Coll Health.* 2020;68:41–51.
40. Duncan MJ, Wunderlich K, Zhao Y, Faulkner G. Walk this way: Validity evidence of iphone health application step count in laboratory and free-living conditions. *J Sports Sci.* 2018;36:1695–1704.
41. Höchsmann C, Knaier R, Eymann J, Hintermann J, Infanger D, Schmidt-Trucksäss A. Validity of activity trackers, smartphones, and phone applications to measure steps in various walking conditions. *Scand J Med Sci Sports.* 2018;28:1818–1827.
42. Presset B, Laurenczy B, Malatesta D, Barral J. Accuracy of a smartphone pedometer application according to different speeds and mobile phone locations in a laboratory context. *J Exerc Sci Fit.* 2018;16:43–48.