STUDY TITLE	Fit24: Using technology to improve activity and sleep in Hispanic youth.
SAP VERSION	2.0
SAP VERSION DATE	May 20 2020

1 Table of Contents

2 Abbreviations

PA	Physical activity
SB	Sedentary behavior
SDT	Self Determination Theory
SAP	Statistical Analysis Plan
BMI	Body mass index

3 Study Aims and Outcomes

3.1 Study Aims

Aim 1. Use a mixed-methods approach (in-depth interviews and surveys) to adapt and refine SDT-informed text message content for Hispanic adolescents.

Aim 2. Test the feasibility of a 12-week goal-setting intervention utilizing a Fitbit and SDT-informed text messaging for improving PA, SB, and sleep, in Hispanic adolescents with obesity.

> **Aim 2 Hypothesis:** This approach will satisfy all a priori feasibility criteria for recruitment, data collection, integrity of the study protocol, internal consistency of measures, technical issues, and satisfaction.

Aim 3. Use in-depth exit interviews to further examine the feasibility and acceptability of the intervention.

3.2 Primary outcome (Aim 2)

Feasibility is the primary outcome and will be evaluated for the development of a larger trial using the following a priori feasibility criteria:[1,2]

1. **Recruitment:** Recruit and enroll 48 Hispanic adolescents 14-16 years of age with obesity from community clinic settings.
2. **Data Collection:** Collect data on 100% of participants at baseline and 85% of participants at post-intervention.
3. **Integrity of Study Protocol:** Participants wear their Fitbit device for ≥ 5 days/week and respond to 80% of text messages when prompted throughout the 12-week intervention.
4. **Internal Consistency of Measures:** Cronbach's alpha for survey scales ≥ 0.7.
5. **Technical Issues:** ≤ 10% technical issues with the transmission and receipt of text messages or Fitbit device functioning.
6. **Satisfaction:** ≥ 80% report 'excellent' – 'good' satisfaction with the intervention.

'Recruitment' and 'Data Collection' criteria will assess the feasibility of recruiting and retaining participants from community clinics.[3] We will record the number of participants that are screened, enrolled, and the time and resources that it takes to complete recruitment activities and reach our recruitment goal (N=48). This data will provide information on the availability of eligible, high-risk youth in partnering clinics and the ability for our recruitment strategies to engage youth from these

settings.[2] This data will also provide preliminary insight on the feasibility of recruiting and disseminating this intervention in community clinic settings in future studies. 'Integrity of Study Protocol' criteria will be used to identify the feasible and acceptable 'dose' of the intervention.[1,3,4] We will use the Fitabase online platform to record the number of days per week in which youth wear the Fitbit device during the intervention. We will also use the Survey Signal text messaging service to record the number of times that each participant responds to text messages when prompted during the intervention. This data will provide information on the feasibility and acceptability of asking youth to wear an activity monitoring device and respond to intervention text messages over a 12-week period. 'Technical Issues' criteria will be used to assess the feasibility of using a Fitbit device and text messaging as behaviour change tools. Delivering the intervention with limited technical issues will ensure that these technology components are capable of delivering behavioural interventions with a high degree of fidelity. 'Internal Consistency of Measures' criteria will assess the appropriateness of our survey instruments for use in a population of Hispanic youth at high-risk for type 2 diabetes.[1,3] Last, 'Satisfaction' will assess acceptability,[3] which is especially important for interventions in minority populations.[5] A survey will be used to assess the degree to which participants found the intervention, including the device, text message content, and the frequency and timing of text messages acceptable.

3.3 Secondary Outcomes (Aim 2)

Physical activity (PA), sedentary behaviors (SB), and sleep (min/day) will be measured in min/day.

3.4 Tertiary Outcomes (Aim 2)

Percentage over the 95th percentile in body mass index (BMI)
Psychological needs (Psychological Need Satisfaction in Exercise Scale), and autonomous motivation (Behavioural Regulation in Exercise Questionnaire-2).

4 Study Design

To address the research question, "**Is a technology-based intervention that promotes PA, SB, and sleep, feasible among Hispanic adolescents with obesity?"** a two-armed, pre-post design will be used to test the feasibility of the intervention compared to a wait-list control group. Limitations of previous studies include the absence of control groups.[6,7] A wait-list control comparison is appropriate for studies in the pre-efficacy phase that aim to assess feasibility.[8] Another limitation of previous studies is the use of self-reported measures to assess health behaviors.[9-11] We will rely solely on research grade accelerometers to objectively measure wake time and sleep behaviors at baseline (T0) and 12- weeks post-intervention. Following baseline assessments, youth will be randomized in a 1:1 ratio to the intervention or wait-list control using a computerized random number generator in SAS version 9.4 (SAS Institute, INC., Cary, NC). After the 12-week study period, youth in the wait-list control group will be invited to participate in the intervention. A subsample of youth (N=15) initially randomized to the intervention will be invited to participate in an exit in-depth interview.

4.1 Randomisation and Blinding

A computer-generated randomization list will be used to assign each subject to the intervention or wait-list control group in a 1:1 ratio. We recognize that differences may exist in PA practices and body weight, T2D morbidity, and intervention treatment effect by sex. However, no formal stratification will be conducted because this is a feasibility study with a limited sample size and the feasibility outcomes are primary. Instead, we will aim to recruit equal numbers of males and females. To minimize selection bias and ensure equal numbers of intervention or wait-list control, randomization will be blocked using random sequences of block sizes (4, 6, and 8). SAS version 9.4 (SAS Institute, INC., Cary, NC) will be used to perform the randomization. To ensure blinding of the study biostatistician, the randomization scheme will be implemented by a staff member at a collaborating community clinic who is not involved in recruitment or daily interactions with study subjects. This individual will be identified when the research team meets with clinic personnel to develop recruitment protocols for each clinic.

5 Statistical Analysis

5.1 General Considerations

Analyses will be conducted on all subjects based on the treatment group assigned (intervention or wait-list control) according to the intent-to-treat principal. All statistical tests will be 2-sided and performed using a 5% significance level, leading to 95% (2-sided) confidence intervals. Variables will be summarized overall and by treatment group intervention, wait-list control). Continuous variables will be described using number of subjects, mean, standard deviation, median, 25th percentile, 75th percentile, minimum, and maximum. Skewness, kurtosis, tests for location and histograms will be examined to assess normality and the need for variable transformations. Bivariate continuous variables will be depicted using scatterplots and the Spearman correlation coefficient. Categorical variables will be described using number of subjects and percentages. Denominators of percentage calculations will be based on the number of subjects in the treatment group (intervention or wait-list control).

5.2 Missing Data

The pattern of missing data will be described for the secondary and tertiary outcomes. Differences in subject characteristics (demographics, BMI) and outcomes by whether or not the subject started the intervention or completed the study will be examined.

5.3 Thematic Content Analysis (Aim 1)

Interviews will be audiotaped and transcribed through a professional transcription service. Transcribed data will be coded by two trained coders using thematic content analysis with a constant comparison approach in a qualitative analysis software program (NVivo Version 12.5, QSR International, 2019).[12] Coders will read each transcript independently. Repeated ideas and important quotes will be identified and used to develop an initial codebook, an organized list of codes and memos. During a second reading of interview transcripts, each interview will be compared and contrasted with the initial codebook. Revisions will be made, and the codebook

will be updated. Through an iterative process, coders will meet to compare codebooks and will revise or reject codes as needed. If disagreements between coders cannot be resolved, a roundtable consisting of co-investigators will be convened to resolve disagreements and make final decisions on codes. All codes will be organized into groups of related concepts that will inform emergent themes.[13] A final reading will determine if all themes are saturated and that there are no additional themes identified. Two key strategies will be used to enhance the rigor and accuracy of the qualitative analysis, we will: (1) use active discourse with respect to coding, organization of coding categories, and identification of emerging themes, and (2) consult with Dr. Deborah Thompson (Co-Investigator), who is a highly trained qualitative researcher who will review all themes and associated quotes, and make suggestions for coding modifications.[14]

5.4 Main Analyses of Secondary and Tertiary Outcomes (Aim 2)

Since this is a feasibility study, quantitative assessment of the outcomes will be considered preliminary. The objective of these preliminary analyses is to demonstrate whether the intervention shows promise in changing the PA, SB, and sleep behaviors. Emphasis will be placed on describing the variability in the outcomes using measures of central tendency including mean, standard deviation, and 95% confidence intervals.

Comparisons of continuous variables between treatment groups will be conducted using t-tests or a non-parametric analogue (Mann-Whitney U test). Comparisons within treatment group will be tested using paired t-tests or the signed rank test. Comparisons of categorical variables between groups will be assessed using Chi-square or Fisher's exact tests; McNeamr test will be used within groups. The reliability of the psychological needs and autonomous motivation scales will be assessed using Cronbach's alpha. Examination of clinically meaningful effects,[15] and estimation of the reliable change index[16] will be conducted.

5.5 Exploratory Analyses of Secondary and Tertiary Outcomes (Aim 2)

Subgroup analyses by gender will be conducted as exploratory using summary statistics since the study is not power to detect an effect by gender.

A repeated measures model will be used to examine change in PA, SB and sleep. Each outcome will be tested in a separate model. The model will include time, treatment group, and the time by treatment group interaction. The magnitude and direction of the associations using parameter estimates and associated confidence intervals. Potential covariates include subject age, gender and BMI.

Compositional data analysis will be conducted for the PA, SB, and sleep behaviors. After log transformation, the geometric mean of each behaviour and all behaviors together (overall) will be calculated. The log-ratio of each behaviour will be derived by dividing the geometric mean of each behavior by the overall geometric mean. Differences by treatment group will be assessed using t-test. Variability of each behaviour in relation to the other behaviors and variance of the whole composition will be described using a variation matrix. The association of treatment group with the behaviors will be examined using multivariate regression (all outcomes as dependent variables in one model), after adjusting for gender.

References

1. Arain M, Campbell MJ, Cooper CL, Lancaster GA. What is a pilot or feasibility study? A review of current practice and editorial policy. *BMC medical research methodology.* 2010;10:67.
2. Kraemer HC, Mintz J, Noda A, Tinklenberg J, Yesavage JA. Caution regarding the use of pilot studies to guide power calculations for study proposals. *Archives of general psychiatry.* 2006;63(5):484-489.
3. Czajkowski SM, Powell LH, Adler N, et al. From ideas to efficacy: The ORBIT model for developing behavioral treatments for chronic diseases. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association.* 2015;34(10):971-982.
4. McVay MA, Bennett GG, Steinberg D, Voils CI. Dose-response research in digital health interventions: Concepts, considerations, and challenges. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association.* 2019;38(12):1168-1174.
5. Thornton PL, Kumanyika SK, Gregg EW, et al. New research directions on disparities in obesity and type 2 diabetes. *Annals of the New York Academy of Sciences.* 2020;1461(1):5-24.
6. Dallery J, Kurti A, Erb P. A new frontier: Integrating behavioral and digital technology to promote health behavior. *The Behavior analyst.* 2015;38(1):19-49.
7. Ridgers ND, McNarry MA, Mackintosh KA. Feasibility and effectiveness of using wearable activity trackers in youth: A systematic review. *JMIR mHealth and uHealth.* 2016;4(4):e129.
8. Freedland KE, King AC, Ambrosius WT, et al. The selection of comparators for randomized controlled trials of health-related behavioral interventions: Recommendations of an NIH expert panel. *Journal of clinical epidemiology.* 2019;110:74-81.
9. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. *The international journal of behavioral nutrition and physical activity.* 2015;12:159.
10. Gomersall SR, Ng N, Burton NW, Pavey TG, Gilson ND, Brown WJ. Estimating physical activity and sedentary behavior in a free-living context: A pragmatic comparison of consumer-based activity trackers and ActiGraph accelerometry. *Journal of medical Internet research.* 2016;18(9):e239.
11. Noah JA, Spierer DK, Gu J, Bronner S. Comparison of steps and energy expenditure assessment in adults of Fitbit Tracker and Ultra to the Actical and indirect calorimetry. *Journal of medical engineering & technology.* 2013;37(7):456-462.
12. Ryan GW, Bernard HR. Techniques to Identify Themes. Field Methods. 2003;15(1):85-109.
13. Bernard HR, Ryan GW. Analyzing qualitative data: Systematic approaches. Thousand Oaks, CA: Sage; 2010.
14. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res. 2005;15(9):1277-88. PubMed PMID: 16204405.
15. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. J Psychiatr Res. 2011;45(5):626-9. PubMed PMID: 21035130; PMCID: PMC3081994.
16. Tingey R, Lambert M, Burlingame G, Hansen N. Assessing clinical significance: proposed extensions to method. Psychother Res. 1996;6(2):109-23. PubMed PMID: 22242565.