COVER PAGE

Official title of the study: Mobile Health Technology for Personalized Tobacco Cessation Support in Laos (SupportLaos)

NCT number: NCT05233228

Date of the document: December 12, 2021 (OUHSC IRB approval date).

STATISTICAL DESIGN AND POWER

R21 Phase: Not applicable.

R33 Phase: Conduct a RCT to evaluate the efficacy of AT

Overview. Participants will be patients recruited through 2 large hospitals in Lao PDR. Consenting participants will be randomized to one of the 2 study groups: Automated Treatment (**AT**; n=250) and Standard Care (**SC**; n=250). Participants will complete in-clinic assessments at baseline, and at 3, 6, and 12 months following study enrollment (*Figure 2*). Participants will also complete brief weekly assessments via smartphone.

Design considerations.

Oversampling female smokers. We acknowledge the importance of ensuring that our mHealth-based AT program works well for women. Thus, in the R21 Phase, we will recruit equivalent numbers of male and female smokers for qualitative interviews/discussions. However, given the limited timeframe and budget of the R33 Phase, we must guarantee a sufficient sample size to demonstrate the efficacy of AT in at least the majority group of smokers (i.e., males) if subgroup analysis is needed. Thus, we cannot oversample female smokers but we will ensure that female smokers are represented in a manner that aligns with smoking prevalence in the real-world. Specifically, because the male to female smoking prevalence ratio in the general population is ~7 (51%:7%),[1] we propose to recruit a sample that includes 15% (n=75) females. By doing so, we recognize that we may not have sufficient power to detect a difference in the primary outcome in the female subsample. However, given that little is known about the factors associated with smoking cessation in Laotian females, this proposal will provide valuable information that may inform future research efforts.

Primary outcome. The primary outcome is smoking status at 12 months post-enrollment. Abstinence will be defined as biochemically confirmed self-reported 7-day point prevalence abstinence with expired CO <6 ppm.[102] Secondary outcomes include 3- and 6-month abstinence. We will consider several other common outcomes, such as continuous and sustained abstinence, number of quit attempts, and length of abstinence.

Statistical analysis. *Hypothesis*: At the 12-month follow-up, 7-day point prevalence biochemically confirmed abstinence (primary outcome) will be higher in the AT (vs. SC) group.

The primary abstinence analysis will be intention-to-treat; patients not completing follow-up assessments will be considered smokers. However, we will explore other ways of dealing with missing data (see below). Based on our previous work and assuming attenuation of abstinence rates across the follow-up visits, we anticipate that the 7-day point prevalence abstinence at the 12-month follow-up will be 8% in the SC group and 17% in the AT group. With 250 participants in each group and assuming a two-sided alpha of 0.05, we will have 86% power to detect a difference of 9% in 12-month 7-day point prevalence abstinence in the overall sample. It is unknown whether factors associated with smoking cessation differ by sex; it seems likely given the variation in smoking prevalence between sexes. Thus, we chose a sample size that gives 80% power to detect this difference in a subgroup analysis of only males (n=425, 85% of the sample). The sub-sample of females (n=75) gives 80% power to detect a 25% difference in 12-month 7-day point prevalence abstinence.

To estimate the effect of AT on abstinence rates while accounting for the potential clustering of participants recruited from the 2 clinic sites, we will use generalized linear mixed models (GLMM) analysis in which intervention groups (AT vs. SC) will be estimated as a fixed effect while the clinic will be modeled as a random effect nested within treatment condition. Sex can also be modeled as a random effect nested within clinic and treatment condition. Specifically, log binomial mixed models will be used to estimate the relative risk of quitting in the AT (vs. SC) group. Although groups should be similar in baseline characteristics due to randomization, we will explore models that control for any demographic or clinical variables that differ between treatment groups at baseline. We will also use GLMM models to examine changes in abstinence rates over time, while accounting for relevant baseline covariates. Similar GLMM or linear mixed model (LMM) methodology, as appropriate for each outcome variable, will be used to examine other smoking-related variables, such as continuous abstinence, prolonged abstinence, and quit attempts. This project focuses on the latter 3 PBM phases, and thus we will explore phase-specific outcomes such as abstinence attainment and number of days smoking/abstinent. Statistical analyses will be performed using SAS 9.4 (SAS Institute, Inc.).

Potential treatment mechanisms will be examined via mediation analyses with intervention group (AT vs. SC) being the independent variable, abstinence at 12 months being the outcome variable, and the hypothesized mechanisms (motivation, self-efficacy, stress/negative affect) being potential mediators. The

PROCESS macro for SPSS/SAS[103,104] will be used to identify variables (e.g., motivation, self-efficacy) that mediate the relationship between treatment condition and smoking cessation outcomes. This method uses an ordinary least squares path analytic framework to estimate direct and indirect effects in single and multiple mediation models, and bootstrapping methods are incorporated to generate confidence intervals.

Missing data and dropouts. Although treating participants lost to follow-up as smoking is a widely used strategy in smoking cessation studies, some researchers point to problems with this approach, especially when comparing treatment arms with differential dropout rates.[105] Thus, we will conduct sensitivity analyses to test for treatment differences assuming different missing data mechanisms. For example, we will consider a multiple imputation approach based on smoking-related patient characteristics at baseline, as well as demographics to account for potential missing-at-random mechanisms. We will also explore pattern-mixture and selection models to account for potential (and likely) missing-not-at-random mechanisms.[106] Similar findings based on these analyses will strengthen our study conclusions.