Research Evaluating Sleep & Trends for Universal Prevention (REST-UP)

NCT04783519

03/27/2024

Statistical Design and Power

Preliminary analysis. Prior to inferential statistics, univariate and bivariate descriptive statistics will be used to assess distributions and simple relations among variables. Preliminary analyses will include nature of missing data, identification of extreme values and variable distributions. To establish randomization effectiveness, baseline equivalence on drinking measures and demographic representation will be examined between conditions. Preliminary efficacy (Aim 1) will be evaluated with respect to drinking behavior (quantity/frequency/peak drinking, total drinks per week, alcohol consequences) and sleep symptoms (as measured by the ISI and Promis). Secondary outcomes include MJ use and consequences.

To examine intervention differences in outcomes from post-intervention through 3-month follow-up, multilevel regression analyses (Gelman & Hill, 2007) – often referred to as Hierarchical Linear Modeling (HLM; Raudenbush & Bryk, 2002) or mixed-effects modeling (Pinheiro & Bates, 2000) – will be used. HLM has a number of advantages over traditional analytic techniques, such as repeated measures ANOVA, in that it allows for missing data, handles varying time-points, accommodates both time-varying and time-invariant covariates, and allows for flexible modeling of the correlation among repeated-measures. In the proposed project, each participant will provide up to 2 repeated measures (post-intervention, 3 months), yielding up to 300 Level 1 cases (repeated-measures) across 150 Level 2 cases (people). The following model will be the basis for analyses of specific aims:

$$\text{Level 1:} \quad DV_{ti} = \pi_{0i} + \pi_{1i}(Time)_{ti} + \varepsilon_{ti} \qquad \varepsilon_{ti} \sim N(0,\ I\sigma_{\varepsilon}^{2}) \qquad \text{(Equation 1)}$$

$$\text{Level 2:} \quad \pi_{0i} = \beta_{00} + \beta_{01}(DV_{pre})_i + \beta_{02}(Tx)_i + r_{00i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11}(DV_{pre})_i + \beta_{12}(Tx)_i + r_{10i}$$

where *t* indexes repeated-measures and *i* indexes participants. DV_{ti} represents the outcome vector for each individual from post-intervention through 3-months. $Time_{ti}$ measures weeks since end of intervention. DV_{pre} is the pre-intervention value of the outcome. (Note that sometimes DV_{pre} is included as the first datum in the outcome vector; most biostatisticians now reserve the outcome in treatment studies for post-randomization data that can be influenced by the treatment, see Harrell, 2004). *Tx* represents contrasts (i.e., dummy codes) for treatment condition. The multilevel model fits subject-specific effects via the random-effects (*r*s in Level 2); these "shrinkage" (Empirical Bayes) estimators are weighted combinations of individual and group level information. The model assumes independence of Level 1 residuals conditional on random-effects, and homoskedasticity and (multivariate) normality of Level 1 residuals and Level 2 random-effects (Atkins, 2005). The model depicted above represents the primary model to test intervention differences in outcome. The necessity of terms in the model will be assessed via Wald statistics for fixed-effects and deviance tests for random-effects. Clinical trials of psychological outcomes assessed repeatedly often have a nonlinear trajectory, which is well-modeled by the natural logarithm of time. If there is a nonlinear trajectory that is not fit via the log transformation, polynomials or spline fits will be used to model the curvature. Alcohol and MJ outcomes are frequently positively skewed, calling into question the use of the Normal distribution as the probability model. Rate and frequency variables, such as number of alcohol problems, often have skewed distributions that are better represented by the Poisson or Negative Binomial distributions (Atkins et al., 2013; Atkins & Gallop, 2007). Depending on observed outcome distributions, a generalized mixed-effects model with log-link a Poisson distribution may be appropriate (Radenbush & Bryk, 2002). This model is a type of HLM with a Poisson distributed Level 1 outcome, the only notable change to equation 1.

Aim 1b: Preliminary Treatment Outcome. The primary comparison is post-intervention differences between the BASICS + SLEEP vs. AOC and BASICS vs. AOC. The coding for *Tx* will depend on which hypothesis is being tested. Coefficients representing the main effect of treatment (i.e., "intercept" differences at post- intervention) and cross-level interactions of treatment contrast and Time will be used to test H1b. These two sets of coefficients will show whether there are treatment differences at the end of therapy and also across the 3-months post-therapy. As noted, appropriate Level 1 distribution (i.e., Normal or Poisson) will be used, determined by empirical distribution of outcomes. Secondary analyses will examine MJ using these approaches. We anticipate small-to-moderate effect sizes for alcohol and MJ use and moderate-to-large effects on sleep symptoms; recent trials have demonstrated large pre-post effects on alcohol and sleep indices for an integrated intervention with college students (Fucito et al., 2017).

Missing data. Multilevel models can incorporate missing data; however, validity of the findings depends on the missingness mechanism (see Atkins, 2005). There are 3 common assumptions for missing data (Schafer & Graham, 2002), and multilevel models provide unbiased results unless data are found to be non-ignorable (NI). HLM is generally resilient to missing data because HLM weights each person's final within-person slope and intercept estimates by the reliability of these estimates, which is influenced by the relative variance components and the number of observations that person provides. Data from individuals with fewer assessment points have less influence on the final results. More problematic is NI data, which arises when missingness depends on unobserved values of the outcome or upon a covariate not included in the model. In the present study, participants could fail to provide data due to alcohol use, which would yield NI missing data. This would only lead to NI data if the missingness were not related to earlier values of the DV.

To deal with this possibility, we will use both design and statistical controls. First, as noted in Approach and Recruitment and Retention Plan in the Human Subjects and Clinical Trials section, we make every attempt to retain all participants at all assessments, and HLM models include all participants in all analyses even if missing data at some assessment points and/or on some items. Second, handling NI missing data is an active area of statistical research, and there are several options that attempt to provide unbiased estimates in the presence of NI missing data, including selection models, pattern-mixture models, and shared parameter models (Schafer & Graham, 2002). We will use the pattern-mixture approach to missing data that has been extended to HLM (Hedeker & Gibbons, 1997) to examine the sensitivity of our primary findings to key patterns of missing data. Should these sensitivity analyses reveal that the findings are dependent on missing data, the pattern-mixture approach will allow us to combine effects across missing data patterns to yield unbiased treatment results (see Hedeker & Gibbons, 1997 and Atkins, 2005 for examples).

Power. We used PASS software to estimate effect sizes based on proposed study sample size for Aim 1b. In Aim 1b, the models will be run such that there will be two repeated measures of the outcome at follow- up. Given 50 participants per condition, this means that there will be up to 200 observations for any one comparison of an intervention to the control group (2 conditions x 50 subjects per condition x 2 observations per subject). However, the effective sample size will be smaller than 200 due to the design effect for the correlation of observations within participants (Snidjers & Bosker, 1999). Thus, we can use an effective sample size to estimate the effect of an intervention compared to control at any given follow-up visit. Given the sample size, 2 repeated observations, and an anticipated intra-class correlation (ICC) among survey responses for the same individual of .6, the effective sample size for any given comparison between two groups would be 125. Using this effective sample size and standardizing the sleep outcome with a mean of 0 and SD of 1, we would expect >.8 power to detect moderate-sized coefficients for the difference between intervention and control as small as .5. If the outcome were typical number of drinks per week, which tends to show a count distribution, and the base-rate were 4 drinks per week, then we expect >.8 power to detect a 32% increase in the count of drinks per week in the control compared to intervention group at any given follow-up (Rate Ratio = 1.32).