Statistical Analysis Plan (SAP) for Feasibility Pilot in Preparation for Large Pragmatic Encouragement Trial of Bright Light Therapy (BLT) for Depression (BLT Pilot/Bright Start Study)

ClinicalTrials.gov identifier	NCT05356130
SAP version	Proposal
SAP version date	June 2, 2021
NIMH FAIN	R34MH122550
IRB of record	Kaiser Permanente interregional IRB

SAP Revision History

Version edited	Date edited	Section number changed	Description and reason for change

INTRODUCTION

This statistical analysis plan provides guidelines for the final presentation and analysis for the BLT Pilot Trial. The R34 pilot was designed to prepare for a future, fully-powered effectiveness trial. This study collects information to conduct the BLT Pilot Trial as described in these following aims:

Aim 1.1: We will enroll 90 depressed patients in this modest-size feasibility pilot, and randomize them to:

- Arm 1: Treatment as Usual (TAU) control: A "usual care" control group (e.g., antidepressants, watchful waiting, psychosocial therapy; all TAU will be recorded for all participants in all study conditions); or
- Arm 2: TAU + Minimal BLT Encouragement: TAU plus two minimal communications (mailed letter, secure EHR message) identifying BLT as a promising treatment, and outlining steps for patients to self-initiate. Arm 2 will not include any phone coaching or adherence promotion.
- Arm 3: TAU + Enhanced BLT Encouragement + Adherence Promotion + MI: TAU plus 2-4 brief calls to encourage BLT use, advise on purchase of a light box (LB), assist with obtaining insurance reimbursement, educate for correct LB use, and provide motivational interviewing (MI) as needed to promote adherence.

Aim 2. This small pilot sample is insufficient for confirmatory tests. However, we will conduct analyses that mirror those planned for a future, fully-powered trial. Key analytic contrasts are (a) active Arms 2 + 3 vs. control Arm 1; (b) each active Arm (2, 3) separately vs. control Arm 1; and (c) Minimal Arm 2 vs. Enhanced Arm 3. For each contrast we will initially examine differences between study arms in pre-post change in proposed mediators in BLT and MI effects; followed by examination of differences in pre-post change in the primary (PHQ-9) and key secondary outcomes (see Assessment). We then examine candidate mediators (improved sleep, normalized circadian rhythm, increased physical activity, readiness for change) for mediating BLT and MI effects on depression. We will also examine moderation of BLT effects and subgroup outcome variation (e.g., receiving vs. not receiving TAU ADs). Pilot feasibility data will facilitate rapid launch of a future trial: estimated recruitment, retention, and adherence with BLT protocol; and a cautious estimate (with wide confidence intervals) of BLT effectiveness.

BACKGROUND

Rationale and research questions

BLT Pilot aims to test the feasibility of conducting a controlled clinical trial designed to evaluate the effectiveness and cost-effectiveness of minimal and enhanced "encouragement" to initiation and regular use BLT for treatment of an incident unipolar depressive episode among adult members of the KPNW medical plan. Encouragement includes motivational interviewing, education, and tools for using BLT with high adherence.

Our primary clinical outcome is the PHQ-9 at post-acute treatment (2 months follow-up). Our primary process outcomes are initiation of BLT therapy and adherence to best BLT practices.

Hypotheses

While this is a pilot study and not powered to assess intervention effects, we hypothesize that the combined intervention (minimal + enhanced encouragement arms) will have larger reductions in PHQ-9 at post-acute treatment than the TAU control arm; and that enhanced encouragement will have larger reductions in PHQ-9 at post-acute treatment than minimal encouragement.

Abbreviations

BLT, bright light therapy; SD, standard deviation; ITT, intention-to-treat; CEA, cost effectiveness analysis; TAU, treatment as usual; LB, light box; PHQ-9, patient health questionnaire 9-item version; HLM, hierarchical linear models; EHR, Electronic Health Record; SE, Standard Error; ISI, insomnia severity index; MI, motivational interviewing; ES, effect size.

Preliminary data evaluation

Before any analyses, we will conduct procedures to assess data quality and completeness and will evaluate missing data patterns. In order to maximize analytic team efficiency, we will audit data from the initiation of data collection. Careful descriptive analyses (e.g., mean, SD, range, skewness) will be conducted for all study variables of interest. We will evaluate distributions to ensure that they meet the assumptions of planned analyses (below), including the detection of outliers (another potential indicator of entry error). As noted early, this proposed R34 pilot is underpowered to perform formal tests of effectiveness. However, we will conduct analyses as we would for a fully-powered, R01-funded project, both to demonstrate feasibility and to generate effect size estimates for the future, fully powered randomized trial—subject to the limits and cautions of using pilot data for this purpose (e.g., use with very wide confidence intervals)[1-3].

Management of missing data

Data can be missing for many reasons, including entry errors, interviewer omissions, treatment refusal, acute medical illness, etc. Missing data is also a consequence of attrition, but we expect dropout from the study will be less than 10% at the 6-month follow-up assessment. To carry out the planned ITT analyses, we use multiple imputation methodology to account for missing data[4, 5]. These methods preserve the correlational structure among the study measures and incorporates the uncertainty inherent in replacing missing values, and as with other imputation procedures, they assume data are "missing at random" (MAR). We will create 10 imputed datasets, repeating outcome analyses across each replicate, and will use Rubin's rules to produce adjusted estimates and statistics from which inferences will be drawn[6].

Analysis of primary outcome

For our PHQ-9 primary outcome, we will test whether there are differences in score trajectories across time between arms with two-level hierarchical linear models (HLMs) in a growth curve framework[7-9]. The first level of the model will include time as a predictor, modeling within-person variation. The second level will include a dummy variable for arm (e.g., 1=TAU+BLT+AP or TAU+BLT, 0=TAU) as the predictor variable for both the random effects of the intercept and slope for time. A significant coefficient for arm on the time slope would indicate that there are different trajectories across time for BLT arms versus TAU. We will probe any significant arm x time interactions by graphing the simple-effects equations to determine whether the observed pattern is consistent with hypotheses. Compared to the Arm 1 TAU condition, a pattern in which the active arms (Arms 2 and 3) exhibit a greater PHQ-9 decrease over time supports the effectiveness of the intervention. However, recall this pilot is insufficiently powered to formally test intervention effects[1-3].

Moderation and mediation

In addition to primary outcome analyses we will conduct exploratory analyses of moderation using variables identified through previous research and clinical expertise, specifically indication of pharmacotherapy use (1=yes, 0=no), age (continuous), gender (1=female, 0=male), race and ethnicity (1=racial/ethnic minority, 0=non-Hispanic white), baseline PHQ-9 (1=clinically elevated, 0=below clinical range), and EHR indication of mental health comorbidities such as eating disorders, anxiety (1=comorbidity, 0=no comorbidity). Moderation will be indicated by a significant time x arm x moderator interaction in HLM.

We propose to conduct exploratory analyses of the BLT mediation pathway as the extent to which the relation between BLT and outcomes is mediated by changes in participants' 1) sleep (sleep diary and actigraphy, and ISI subjective sleep quality), 2) total physical activity, and 3) circadian timing and chronotype. For example, to evaluate the mediating effect of ISI sleep quality we will regress post-acute treatment ISI on baseline ISI and a randomization condition indicator; and we will regress follow-up PHQ-9 at post-acute treatment ISI and a randomization condition indicator. Using the product coefficients mediation approach, we will use the resulting regression coefficients and standard errors (SEs) to estimate the indirect effect[10-12]. Because SEs of the products of regression coefficients are not normal, we will use bias-corrected bootstrapping to estimate the SEs for the indirect effect.[12] We will use a similar approach for the MI pathway to test the mediating effect of participants' increased readiness to change on BLT adherence.

Analysis of secondary outcomes

Continuous measures over time will be modeled with HLM methods described above. Health service use can be categorized into binary (e.g., met threshold for medication possession ratio, had > 1 hospitalization) and count outcomes (e.g., N of ED visits, primary care visits). Binary outcomes will be modeled with multivariable logistic regression, and the coefficient of study arm membership (e.g, 0=TAU, 1= TAU+BLT+AP & TAU+BLT) is the coefficient of primary interest. Count outcomes are modeled with either multivariable Poisson or negative binomial regression depending on which yields optional fit.

BLT process outcomes

To confirm that the Arm 3 adherence/MI elements had their intended effect we will contrast Arms 2 vs. 3 across months 2, 4, and 6, comparing: (a) % cumulative light box purchase; (b) % cumulative DME reimbursement; and (c) mean EMA-adherence (aka, dose). We will conduct these analyses using methods similar to those for secondary outcomes; however, the analysis sample will be restricted to Arms 2 and 3. We will describe similar metrics in the TAU condition (Arm 1) but expect little BLT uptake there.

Economic analysis

The future full trial will conduct a cost-effectiveness analysis (CEA) of the BLT intervention relative to TAU to inform the merits and costs of implementing such a program. In this pilot we will conduct both the relevant data collection and the economic analyses to prepare for this. Two major CEA cost components include 1) TAU costs related underlie all study arms, and 2) costs of the BLT intervention, including the equipment but also phone coach staff time[13]. While this pilot is underpowered to detect significant difference in costs between study conditions, we will be able to richly describe and

categorize TAU services underlying each condition. We will deploy data collection tools for phone coaches and research staff to track time used to deliver the interventions, and will periodically meet with those who are providing data to assess whether the tools are meeting their needs and if the data are truly capturing the resource intensity needed to deliver the protocol. This process will help inform the future full-scale trial about the types and patterns of TAU services as well as the best way to collect and value the resources needed to deliver this type of intervention.

Power analysis, sample size

The R34 pilot funding cap limits us to a feasibility sample that is insufficient for an adequately powered test of the hypothesis. Nonetheless, we provide an estimate of the statistical power obtained from this sample of 90 participants. Meta-analyses suggest an overall effect size for BLT of d=0.84[14]. We conservatively estimate the effect size likely to be obtained in this pilot trial as lower than reported in meta-analyses because this is a real-world effectiveness population; even with Arm 3 adherence promotion BLT adherence is likely to be lower than in efficacy trials, leading to lower clinical benefit. Assuming a retained sample size of N=81 at the final 6-month assessment, we conducted power analysis for the HLM of the primary outcome (using SAS version 9.4) to determine the minimum detectable effect size. At a two-tailed alpha level of .05 and conservatively assuming an autocorrelation of .70, we would have 80% power to detect a Cohen's d for the time by arm effect as small as 0.39 for the contrast between active intervention arms (Arms 2 and 3) and TAU, and as small as 0.55 for contrasts between individual arms; both estimates are medium effects.

Use of pilot results to estimate power for the future full trial

Regardless of the between-condition effect size that we observe in this pilot, we will heed the cautions that have been raised about using pilot study results to estimate effect for full trial power/sample calculations[1-3]. We will consider between-condition effect size (ES) results highly tentative given the small pilot sample, and will use it only with wide confidence intervals when employing it to estimate power for the future full pragmatic trial. For example, if we obtain an effect size of d=.45, we will generate a power/sample size matrix with estimated effects of .35, .40, .45, .50—all ranging around the observed effect from this pilot, but trending to somewhat smaller and more conservative effects.

References

1. Leon, A.C., L.L. Davis, and H.C. Kraemer, The role and interpretation of pilot studies in clinical research. J Psychiatr Res, 2011. 45(5): p. 626-9.
2. Tickle-Degnen, L., Nuts and bolts of conducting feasibility studies. Am J Occup Ther, 2013. 67(2): p. 171-6.
3. Kraemer, H.C., et al., Caution regarding the use of pilot studies to guide power calculations for study proposals. Arch Gen Psychiatry, 2006. 63(5): p. 484-9.
4. Enders, C., Applied Missing Data Analysis. 2010: Guilford Press.
5. Little RJ, R.D., Statistical Analysis with Missing Data. 2 ed. 2002, New York: John Wiley & Sons.
6. Rubin, D., Multiple Imputations for Nonresponse in Surveys. 1987, New York: Wiley.

7. Hox, J., Multilevel analysis: techniques and applications. 2010: Routledge.
8. Bryk AS, R.S., Hierarchical linear models: Applications and data analysis methods. 2 ed. 2002, Thousand Oaks, CA: Sage Publications, Inc.
9. Snijders T, B.R., Hierarchical linear models: Applications and data analysis methods. 2 ed. 2002, Thousand Oaks, CA: Sage Publications, Inc.
10. Hayes, A.F., Beyond Baron and Kenny: Statistical Mediation Analysis in the New Millennium. Communication Monographs, 2009. 76(4): p. 408-420.
11. MacKinnon, D., Introduction to Statistical Mediation Analysis. 2008: Routledge.
12. MacKinnon, D.P., et al., A comparison of methods to test mediation and other intervening variable effects. Psychol Methods, 2002. 7(1): p. 83-104.
13. Gold MR, S.J., Russell LB, Weinstein MC, Cost-Effectiveness in Health and Medicine. 1996, Oxford: Oxford University Press.
14. Maruani, J. and P.A. Geoffroy, Bright Light as a Personalized Precision Treatment of Mood Disorders. Front Psychiatry, 2019. 10: p. 85.