

Statistical Analysis Plan

Protocol Number: CIRB 20-08

Work as a determinant of health: A pragmatic trial of enhanced cognitive behavioral therapy to competitive work and wellness in veterans with serious mental illness (WORKWELL)

Funding Agency: HSR&D

9/30/2025

Analysis Plan

Statistical Considerations

Sample size and power

276 total participants will be needed to ensure > 80% power for this two-arm trial, accounting for the clustering effect of CBTw group cohort and potential attrition of 13%.

Aim 1: Work outcomes. The primary analysis is intention-to-treat and the primary outcome is total weeks worked in competitive jobs with primary endpoints of 6 and 9 months. Power estimate for total weeks worked was calculated using simulated data from a three-level linear regression model to account for the nested structure of the data. Simulation studies are the most flexible method for power analysis which can be tailored to any kinds of analyses. SAS 9.4 was used to generate 1000 datasets from each hypothesized model with plausible effect sizes for parameters of interest based on pilot data. Power estimates are based on the percentage of the datasets (replications) in which the parameters were significant. Based on the open trial pilot data, an ICC of .077 was estimated for the clustering effect of cohort and an ICC of .2 for repeated measures. Our pilot data show that CBTw significantly increased total weeks worked at 6 months ($f^2 = .15$). It is anticipated that there will be a larger increase in total weeks worked over time for the treatment arm. With the target sample size, there will be 97.6% power to detect even a small study arm by time interaction effect ($f^2 = .03$) with type I error set at .05. Further, the outcome distribution is likely to be non-normal. Although power will be attenuated by this non-normality, power to detect the target effects is still expected to be sufficient (>80%). For the binary outcome of achievement of steady competitive work, CBTw pilot data show that 49% of participants who were unemployed at baseline became steady workers at 6 months. In WORKWELL, it is expected that 5-10% more participants will acquire work and become steady workers during by the 9 month follow-up. Thus, power estimates may be conservative. In the control arm, it is anticipated that approximately 34% will become competitively employed and a lesser number will be steady workers by 9 months, based on VHA program evaluation data reporting 1-year SE outcomes. Based on the effect size, simulated data was used from two level logistic regression models, finding that the target sample size will provide > 99% power to detect the treatment effect, accounting for a conservative estimate of 13% attrition. If more participants gain competitive employed in the control arm (~50%), there will be >80% power to detect this effect.

Aim 2: Health and recovery outcomes. Two-level linear regression models accounting for the nested nature of the data (repeated measures nested within individuals) will be used to examine the effects of CBTw + SE services compared to psychoeducation + SE on the change of each of the continuous outcomes including recovery (RAS), health-related quality of life (PCS & MCS), symptoms (DSM), and suicidality (SSI) over time (baseline and 3-, 6-, 9-months follow-ups). It is expected that there will be adequate power (>80%) to detect the treatment effect on health and recovery outcomes even if the effect size is small (i.e., $f^2 = .03$). *Regarding the outcome of suicidal ideation, it is anticipated that the range of responses will be smaller, given that fewer participants will report these thoughts (21.2% of participants in the CBTw open pilot reported passive ideation at baseline; 9.6% reported such thoughts at follow up). Thus, the study may be less powered to detect a treatment effect. However, despite this issue of range restriction given the relatively lower base rate of suicidal thoughts, WORKWELL will nonetheless examine possible treatment effects, due to the importance of this outcome in Veterans with SMI.*

Preliminary analysis: Baseline characteristics of sample & covariates analyses

Given the RCT design, characteristics of participants are expected to be balanced across the two study arms. Baseline characteristics and baseline clinical outcomes (e.g., symptoms) will be compared between the arms using the appropriate test (i.e., t-test, or Chi-square test). Variables found to significantly differ will be included as covariates in the primary analyses. All primary analyses will also include time receiving SE services and site membership as covariates. The potential interaction effects of site with key predictors will also be included as predictors to account for the site effect. Lastly, attendance (the number of CBTw or psychoeducation sessions attended) and the association with the Aim 1 and Aim 2 outcomes will be examined.

Aim 1 Analyses: Competitive Work Outcomes

Hypothesis 1: For the outcome, total weeks worked in competitive jobs, three-level linear regression models will be used to test the effects of CBTw + SE compared to psychoeducation + SE over time, with predictors of study arm, arm by time interaction, site, and interaction between site and key predictors, covariates, as well as random effects to accommodate the potential non-independence due to repeated measures and group cohort effects. The time effects could be linear or nonlinear; these possibilities will be examined by treating time as continuous or categorical. If the outcomes are not normally distributed, robust maximum likelihood estimation method will be used to correct the standard error estimates for nonnormality. In the case of excess of zeros for this outcome, hurdle negative binomial models, a reliable method to analyze zero inflated longitudinal data of this nature, will be utilized. For the binary outcome, steady competitive work status across the 9-month study period, two-level logistic regression model will be used with the same set of predictors as in the linear models, with random effects to accommodate the potential non-independence due to group cohort effects.

Aim 2 Analyses: Health and Recovery Outcomes

To test **Hypothesis 2a**, Two-level linear regression models accounting for the nested nature of the data (repeated measures nested within individuals) will be used to examine the effects of CBTw + SE services compared to psychoeducation + SE on the change of each of the continuous outcomes including recovery (RAS), health-related quality of life (PCS & MCS), symptoms (DSM), and suicidality (SSI) over time (baseline and 3-, 6-, 9-months follow-ups). The scores of the outcome measures are calculated by averaging the corresponding items. The models are the same as those in Aim 1 analysis except that the outcomes are different. It is hypothesized that CBTw + SE will lead to a greater improvement on the health and recovery outcomes (indicated by a significant arm by time effect). If the outcomes are not normally distributed, robust maximum likelihood estimation method will be used to correct the standard error estimates for nonnormality.

To test **Hypothesis 2b**, two-level linear regression models will examine the effects of work on health and recovery outcomes over time, with predictors of competitive weeks worked, time, weeks by time interaction, site membership and potential interaction between site and key predictors, potential covariates (i.e. controlling for baseline levels of health and recovery), as well as random effects to accommodate the potential non-independence due to repeated measures and group cohort effects.

Missing data: Aims 1 & 2

Missing data from Aims 1 and 2 will come in two different forms: missing by attrition and intermittent missing of observation. Using logistic regression, the effects of missing observations due to sample attrition will be examined by investigating the participant characteristics associated with dropout. If results reveal that data are missing completely at random (MCAR) or due to observed factors (MAR), the data will be analyzed using full information maximum likelihood, taking into account potential predictors for missingness. However, if the pattern of missing data is nonignorable, pattern-mixture models that includes variables defined by the participant's pattern of missing data will be used. Sensitivity analyses also will be performed to examine how the result will change based on different assumptions about the missing data, such as the missing work period having no work, the mean amount of work as available data from earlier follow-up periods, or full employment.

Account for Heterogeneity in the Sample

The study is limited to working aged Veterans with SMI who are unemployed at baseline, decreasing the heterogeneity of the sample. In addition, analyses will examine whether treatment effect will change as a function of the site membership and major characteristics of the study participants, such as different types of SMI diagnoses, age, and current substance use (e.g., users vs. non-users).

AIM 3 CBTw IMPLEMENTATION: OVERVIEW

The initial CBTw implementation strategy was informed by implementation science and developed based on our experience in the two CBTw pilots, stakeholder surveys and interviews, and feedback from operational partners. *In order to ensure successful implementation of CBTw during WORKWELL and wide-scale uptake following its completion, an evaluation of this initial implementation strategy and barriers and facilitators of implementation will be conducted, guided by the RE-AIM framework.*

Initial Implementation Strategy

As identified in the ERIC guidelines and based on the procedures established in CBTw pilots and our experience implementing other psychosocial interventions.

- (a) **Implementation planning** (Months 1-2): During initial start-up, the study team will meet with SE staff and managers at sites to determine the best approach to integrating CBTw in work flow, and develop procedures to support implementation (e.g., documentation procedures). The study team will also work with sites to select intervention facilitators chosen from the VRCs at each SE program. **VRCs will deliver the CBTw and psychoeducation interventions (separate VRCs for each condition).** VRCs were chosen because they are present at most SE programs nationally (high generalizability potential) and have master's level training in both employment services and rehabilitation counseling. This implementation approach was suggested by Dr. Bakken (VACO TSES Director) and endorsed by SE managers as appropriate and feasible.
- (b) **Training** (Months 1-3): Kukla, McGuire, Strasburger will conduct onsite training. VRCs delivering CBTw will attend a half-day CBTw workshop consisting of: 1) Education on CBTw content and application of cognitive, behavioral, and narrative strategies; 2) Mock practice of facilitation strategies; 3) Customization of CBTw content to meet unique needs of Veterans. Second, led by Dr. McGuire, psychoeducation facilitator training will follow a parallel structure. Third, other SE staff will be oriented to study procedures. Fourth, SE managers overseeing VRCs will receive training on how to provide supervision to ensure CBTw implementation.
- (c) **Fidelity monitoring & feedback** (~Monthly): Lead by Dr. Rollins, adherence to the CBT model will be assessed by the Cognitive Therapy Scale-Revised (CTS-R). All CBTw and psychoeducation sessions will be audio-recorded to rate fidelity. *At each site, 3 sessions from each group cohort (i.e. 1 session out of every 4) will be randomly selected for rating (in the CBTw and psychoeducation arms).* On a monthly basis, consultation lead by Kukla and Rollins, will include review of fidelity scores and recommendations to improve as needed. *Psychoeducation facilitators will also be given fidelity feedback if they begin to drift toward CBT elements.*
- (d) **CBTw consultation** (weekly) will be led by Kukla and Strasburger via tele or video with VRC CBTw facilitators and SE program managers. Consultation will feature the CBTw Facilitator Guide (Appendix 3) that was piloted in earlier work. In addition, strength based consultation will focus on Veteran case conceptualization, with an emphasis on understanding participants' recovery paths, applying CBTw strategies to fit goals and employment trajectories as they evolve over time, and problem solving obstacles.

Aim 3 RE-AIM Evaluation: Measures & Procedures guided by the RE-AIM Framework

- (a) **RE-AIM domains of implementation, adoption, and maintenance** will be examined through semi-structured **key stakeholders' interviews**. The interviews will also focus on understanding reasons to uptake and continue CBTw and important issues related to inner context and outer context (See Appendix 6). Interviews will also centrally focus on identifying notable
- (b) **RE-AIM implementation, the primary Aim 3 outcome**, will further be gauged using **fidelity assessment**. Fidelity will be measured by the CTS-R, comprised of 12 Likert items (scale of 0-6), with higher scores indicating greater adherence to the CBT model. Acceptable fidelity is a score of 4 or higher, indicating item "proficiency". Used extensively, including in the CBTw open pilot, the CTS-R has strong reliability.
- (c) **Barriers and facilitators to CBTw implementation** within SE programs will be examined through stakeholder interviews with (Appendix 6), and further through virtual site visits (See Section E2d below).

- (d) **Virtual site visits** at pre-implementation (1 month), early implementation (~Month 6), mid-implementation (first quarter-Year 2), and later implementation (first quarter-Year 3) will be conducted. During day-long visits, Drs. Kukla and McGuire will interview VRCs delivering CBTw, SE staff and managers, and make observations of CBTw processes. The implementation support team will keep detailed field notes recording activities conducted, **barriers and facilitators** encountered, and **strategies utilized to address these**. Site visitors will produce a written report that will provide a narrative summary of implementation strengths and weaknesses.
- (e) **RE-AIM Reach** will be examined using **quantitative data**: *CBTw and psychoeducation completion, session attendance, study recruitment, attrition rates; these data will be collected and compared across arms.*

Aim 3 Analyses & Integration of Mixed Methods Data

A deductive-inductive rapid qualitative analysis process will be used to facilitate the timely use of qualitative data to refine the implementation strategy, an approach appropriate for the description of implementation processes and environments. ***Templated summaries*** of each data collection episode will be created, organizing quotes or observations by pre-established codes (the deductive component) pertaining to RE-AIM domains. Emergent themes in the data (the inductive component) will be captured through use of a grounded theory approach, allowing for discovery of themes not included in *a priori* codes. Field notes will be written after **site visit observations** and **staff interviews** will bolster data not captured in immediate data summaries. An iterative process will occur, in which analysis will begin at the first observation and will inform direction of future data collection. The team will then create a **matrix summarizing data by code and collection episode**. *This will facilitate examination of variation across respondents, identify data gaps, and develop summaries of data mapping onto RE-AIM domains, as well as barriers and facilitators encountered.*

Integration of Quantitative and Qualitative Data: **1)** Descriptive fidelity data and **2)** RE-AIM reach data will be calculated for each site. These data will be integrated with **3)** stakeholder interview and **4)** site visit data described above. **Integrated site summaries** will evolve across the study period and will facilitate on-going refinement and finalization of a CBTw implementation strategy bundle.