

Document Type: Statistical Analysis Plan

Document Date: April 6, 2021

**Outcomes Table**

Type of Outcome	Description of Measure	Scale or Score	Time period
Primary	<i>Clinical Assessment Interview for Negative Symptoms</i> (CAINS, Kring et al., 2013), a 13-item measure of negative symptoms that includes two subscales: expression (4 items) and motivation and pleasure (9 items).	CAINS Motivation and Pleasure Scale (CAINS-MAP)	Baseline-12 weeks
Primary	<i>Clinical Assessment Interview for Negative Symptoms</i> (CAINS, Kring et al., 2013), a 13-item measure of negative symptoms that includes two subscales: expression (4 items) and motivation and pleasure (9 items).	CAINS Motivation and Pleasure Scale (CAINS-MAP)	Baseline-24 weeks
Primary	<i>Social Functioning Scale</i> (SFS, Birchwood et al., 1990) measures engagement in social activities through assessment of seven domains: Social Engagement/Withdrawal, Interpersonal Communication, Prosocial, Independence-Performance, Recreation, Independence-Competence, and Employment. In the present analysis, we used a total score for the SFS including all subscales except the employment subscale.	Social Functioning Scale	Baseline-12 weeks
Primary	<i>Social Functioning Scale</i> (SFS, Birchwood et al., 1990) measures engagement in social activities through assessment of seven domains: Social Engagement/Withdrawal, Interpersonal Communication, Prosocial, Independence-Performance, Recreation, Independence-Competence, and Employment. In the present analysis, we used a total score for the SFS including all subscales except the employment subscale.	Social Functioning Scale	Baseline-24 weeks
Secondary	<i>Clinical Assessment Interview for Negative Symptoms</i> (CAINS, Kring et al., 2013), a 13-item measure of negative symptoms that includes two subscales: expression (4 items) and motivation and pleasure (9 items).	CAINS Social Motivation and Pleasure Scale (CAINS-MAP)	Baseline-12 weeks
Secondary	<i>The Role Functioning Scale</i> (RFS, Goodman et al., 1993) is a reliable and valid interview that assesses four domains: Working Productivity, Independent Living/Self Care, Family Network Relationships, and Immediate Social Network Relationships (e.g., Edmondson et al., 2012, Yamada et al., 2010). In the present analysis, we summed ratings of Independent Living/Self Care, Family Network Relationships, and Immediate Social Network Relationships as our measure of role functioning. We excluded Working Productivity because our sample was comprised largely of adults who were retired or not working.	Role Functioning Scale	Baseline-12 weeks
Secondary	<i>Brief University of California San Diego Performance-based Skills Assessment</i> (UPSA; Mausbach et al., 2007) assesses the respondent's capacity to perform tasks similar to those encountered in daily life in two domains: 1) financial skills, in which participants are required to count change, make change from an item purchased at a store, and write a check for a utility bill, and 2) communication skills, in which participants are asked to demonstrate how to use a telephone to dial emergency services, call information to ask for a telephone number, and call a physician to reschedule a medical appointment. For each domain, the total percent correct is calculated and converted to a standardized score ranging from 0 to 50. A summary score is then calculated by summing the two domain scores (range = 0–100), with higher scores indicating better functional capacity.	UPSA-B Total Score	Baseline-12 weeks
Secondary	The 25-item <i>Maryland Assessment of Recovery Survey</i> (MARS; Drapalski et al., 2012) assesses personal recovery in serious mental illness with items addressing having control over one's life, strength to meet and solve challenges, and belief in one's abilities (e.g., "I have abilities that can help me reach my goals", "I am able to set my own goals in life", "I can influence important issues in my life").	MARS Total Score	Baseline-12 weeks

Secondary	<i>Clinical Assessment Interview for Negative Symptoms</i> (CAINS, Kring et al., 2013), a 13-item measure of negative symptoms that includes two subscales: expression (4 items) and motivation and pleasure (9 items).	CAINS Social Motivation and Pleasure Scale (CAINS-MAP)	Baseline-24 weeks
Secondary	The <i>Role Functioning Scale</i> (RFS, Goodman et al., 1993) is a reliable and valid interview that assesses four domains: Working Productivity, Independent Living/Self Care, Family Network Relationships, and Immediate Social Network Relationships (e.g., Edmondson et al., 2012, Yamada et al., 2010). In the present analysis, we summed ratings of Independent Living/Self Care, Family Network Relationships, and Immediate Social Network Relationships as our measure of role functioning. We excluded Working Productivity because our sample was comprised largely of adults who were retired or not working.	Role Functioning Scale	Baseline-24 weeks
Secondary	<i>Brief University of California San Diego Performance-based Skills Assessment</i> (UPSA; Mausbach et al., 2007) assesses the respondent's capacity to perform tasks similar to those encountered in daily life in two domains: 1) financial skills, in which participants are required to count change, make change from an item purchased at a store, and write a check for a utility bill, and 2) communication skills, in which participants are asked to demonstrate how to use a telephone to dial emergency services, call information to ask for a telephone number, and call a physician to reschedule a medical appointment. For each domain, the total percent correct is calculated and converted to a standardized score ranging from 0 to 50. A summary score is then calculated by summing the two domain scores (range = 0–100), with higher scores indicating better functional capacity.	UPSA-B Total Score	Baseline-24 weeks
Secondary	The 25-item <i>Maryland Assessment of Recovery Survey</i> (MARS; Drapalski et al., 2012) assesses personal recovery in serious mental illness with items addressing having control over one's life, strength to meet and solve challenges, and belief in one's abilities (e.g., "I have abilities that can help me reach my goals", "I am able to set my own goals in life", "I can influence important issues in my life").	MARS Total Score	Baseline-24 weeks

## Statistical Analysis Plan

**Statistical Analysis Overview:** Proposed sample size of  $N = 108$  was chosen to detect a standardized mean difference in change from baseline to post-treatment equal to 0.50 between the intervention and active control arms. There were two primary outcome scales – one measuring social functioning and the other measuring negative symptoms. Threshold for statistical significance was  $p < 0.05$  with no adjustment for multiple comparisons.

**Statistical Test of Hypothesis:** The analysis of the primary outcomes utilized a mixed effects model accounting for a random therapy group effects and the analysis included all available data at baseline, post-treatment, and follow-up time-points. Error terms across time points were assumed correlated with unstructured correlation matrix. To test time-specific hypotheses, terms in the mean model included separate indicators for post-treatment and follow-up time points, and interaction terms between these indicators and condition. Based on randomization to treatment condition, the model assumed no differences at baseline, i.e. any differences were assumed to be due to random sampling error. All continuous outcome scales were checked for skew, and an appropriate transformation was applied as needed. Statistical significance was defined as  $p \leq .05$ . All analyses were performed using SAS version 9.4. In moderator analyses we examined whether treatment effects on the primary outcomes (CAINS MAP, SFS) differed according to baseline dysfunctional attitudes measured by the DAS and SARA-Q. We also examined whether baseline negative symptoms (CAINS MAP) modified treatment effects for the SFS. Each moderator variable was dichotomized at the median. In separate models, the moderator and interaction terms with the moderator were added to the primary linear mixed effects models to test whether treatment effects at post-treatment and follow-up time points differed for those above versus below the median.