Primary analysis will be conducted between the placebo group and the tocilizumab group, and the primary outcome will be the PANSS Total scores. This measure has been well-validated in prior clinical studies and is expected to show standard normal distribution, although this will be verified prior to data analysis. Comparisons on other behavioral, cognitive, functional, and biochemical measures will be secondary outcomes. All subjects who have participated in a post-randomization rating will be included in the analysis.

As an initial analysis, we will investigate group differences in PANSS Total scores by using a two-group t test, comparing the mean change across the treatment period (i.e., mean change at 4 week, 8 week, and 12 week time points). If this analysis suggests an effect, we will perform repeated measures analyses (ANOVA) to assess for effects of time (i.e., group by repeated measure interaction) on the primary outcome. p values will be set at 0.05.

Anti-cholinergic agents will be considered covariates in our cognitive analyses. Pearson product moment or spearman rho correlations, depending on observed distributional properties, will be performed on change scores to determine interrelationships in improvement between clinical/cognitive measures, cytokine levels, and tocilizumab doses. If there are major biases within or between treatment groups, effort will be made to assess the impact on results (99).

Sample size calculations are conservatively based upon expectation of a moderate effect-size improvement (d=0.6) in the PANSS Total score in tocilizumab group compared to placebo group. Based upon a sample size of 45 in the tocilizumab group and 45 in the placebo group, the study will have power of 0.8 to detect a moderate effect size (d=0.6, with α < 0.05) for treatment related change in the PANSS total score for the tocilizumab group compared to the placebo group. Other measures are experimental, so power to detect change cannot be computed. Nevertheless, pilot data of this type is essential for establishing feasibility of these interventions, determining effect sizes and competing successfully for large scale funding.