

The Role of Vasopressin in the Social Deficits of Autism

NCT01962870

January 11, 2019

Antonio Hardan, MD, Principal Investigator  
Karen Parker, PhD, Principal Investigator  
Stanford University  
Stanford, California 94305

## Statistical Analysis Plan

Data will be managed using REDCap and analyzed using Least-Squares General Linear Models (LS-GLMs) in JMP Pro 13 and SAS 9.4 for Windows (SAS Institute Inc., Cary, NC). Efficacy analyses will be guided by our pre-study aims which included: 1) testing for treatment main effects on the primary and secondary outcome measures; and 2) testing whether pretreatment blood-based biological measures predicted who benefited from treatment. To minimize the risk of false discovery, we will first identify a robust model using our primary outcome measure (i.e., change in the SRS-2 *T*-Score), and then apply that model to all other outcomes.  $P < 0.05$  will be considered significant and two-tailed tests will be used throughout. Post-hoc tests will be Bonferroni corrected to maintain a family level of  $P < 0.05$  (as detailed below).

The initial model will include sex, ethnicity, weight, IQ, and blood collection time as blocking (control) factors. Pretreatment SRS-2 *T*-Score will be included as a blocking factor to account for the range of possible improvement, and thus reduce possible floor or ceiling effects. Finally, we will include treatment condition (i.e., AVP or Placebo) to test our main hypothesis, and pretreatment blood AVP concentration and pretreatment blood neuropeptide receptor gene expression (expressed as an AVPR1A-to-OXTR ratio to account for within-individual differences in expression) as measures of endogenous AVP function that might predict treatment efficacy. As we administered two different AVP doses according to age, dose will be nested within treatment condition to explicitly test for an overall effect of AVP treatment, and secondarily to control for any dose-related effects. We will also test for treatment condition  $\times$  biological measure and dose  $\times$  biological measure interactions, as pretreatment biological measures should generally only predict treatment outcome in the drug-treated individuals. Finally, because the SRS is a parent reported measure, we will collect SRS scores at two pretreatment time points to identify the reliability of an individual participant's scores. This will enable us to use Weighted Least-Squares General Linear Model (WLS-GLM) analyses whenever parent-reported measures are assessed. WLS ideally uses the inverse of the variance of a mean estimate as the weight, which we will be able to obtain directly from our two pretreatment SRS scores.

Non-significant interactions will be removed in order to avoid confounds of marginality for the main effects, and to distinguish blocking factors. We will test dose nested within treatment condition and pretreatment blood AVP concentration as main effects, and the interaction of pretreatment blood AVP concentration with dose-nested-within-treatment condition. Thus, the model will contain the biologically and experimentally essential variables, regardless of significance. We will further test the robustness of this model by confirming that key results will be held when different blocking factors are included or excluded and if WLS-GLM is not employed. Effect sizes will be calculated as  $\eta^2$  (partial eta-squared), as appropriate for complex linear models. Equivalent Cohen's *d* will be provided for main effects where justifiable. The same model will be applied to all secondary outcome measures with the exception that the baseline, pretreatment behavioral measure will be replaced to match the outcome variable. For child performance measures, LS-GLM, not WLS-GLM, will be used. For CGI-I, in which there is no baseline pretreatment measure, the model will not include a baseline control. Effect sizes for these secondary outcome

measures will likewise be calculated as  $\eta^2$ .

To minimize the risk of false discovery from multiplicity, we will first test the total score for each instrument, and then only test subscales if the total scores are significant. Subscales will be Bonferroni-corrected for multiple comparisons. The assumptions of WLS-GLM (linearity, homogeneity of variance, and normality of error) will be confirmed *post-hoc* and suitable transformations will be applied as needed. *Post-hoc tests* will be performed as planned contrasts and further Bonferroni-corrected for multiple comparisons. We will test our first aim by testing for an overall effect of treatment condition at the mean pretreatment blood AVP concentration.

Fisher's exact Test will be used to test for differences between treatment conditions for participant characteristics and concomitant medications.  $\chi^2$  Likelihood Ratio will be used to test whether parents are able to accurately ascertain the treatment condition to which their child has been randomized. Pearson's correlation coefficient will be used to test the relationship between change from baseline in parent-reported SRS *T*-Score and clinician evaluated CGI-I Score in AVP treated participants. One-way LS-GLMs will be used to test the change from baseline between the AVP and Placebo treatment groups following 4-week treatment for adverse events and safety measures, as well as to test the difference in post-treatment bottle weights between treatment conditions. Suitable transformations will be applied as needed.