Study Title: Patient-Centered Care for Opioid Use Disorders in Federally Qualified Healthcare Centers and Specialty Care Settings

NCT number: Not yet available

Document Date: 11/21/17

Data Analysis Plan

Prior to performing analyses, we will apply standard data screening and cleaning procedures including reviewing the data for data-entry errors, assessing the extent and pattern of missing data, and checking that appropriate assumptions of normality are met. We will employ remedial measures such as power transformations when necessary. We will check that the treatment conditions do not differ at baseline on relevant background variables using analyses of variance (ANOVA) for continuous variables and log-linear models for discrete or ordinal responses. Randomization should minimize the need for covariates to reduce bias in the treatment comparisons, and the primary (i.e., full sample) analyses will not include covariates. In subgroup analyses, relevant covariates will be considered for inclusion as a sensitivity check for randomization and to improve the precision of estimation of the effects [1]. The hypotheses and statistical procedures to analyze the data are listed below. Analyses will be performed using SAS software version 9.1.3.

Primary Hypothesis 1: Clients in the PATH condition will demonstrate higher rates of confirmed substance abstinence (hair and UA verified) during treatment and at follow-up than clients in the SC condition. Confirmed abstinence is the "gold standard" of direct clinical outcomes in SUD research. A mixed effects logistic regression model [2,3] will be used to compare the binary outcome of hair and urinalysis-confirmed abstinence in the two treatment conditions at months 3, 6, 9, 12, 15, and 18. Mixed effects models have advantages of traditional repeated measures approaches in that they allow for missing observations, accommodate measurements made at different time points, provide flexibility in modeling the variance-covariance matrix, and permit the estimation of both group and random patient-specific effects. The model will include terms for treatment condition, time, and their interaction, and the binary baseline abstinence variable along with any other necessary covariates. We will follow the recommendations of Varadhan and Seeger [4] to test heterogeneity of treatment effect (HTE). We have identified two primary subgroup analyses to be performed: (1) setting [IP (n = 400) vs. OP (n = 400)] and (2) type of opioid [heroin (n = 534) vs. prescription (n = 266)]. We will use similar mixed effects logistic regression models (described above) that will additionally include terms for the specific subgroup variable and its interaction with treatment condition. In the event that the interaction is significant, the treatment effect will be estimated separately at each level of the categorical variable that is used to define the subgroups. In addition, we will conduct an *exploratory* subgroup analysis for medication type (agonist vs. antagonist/drug free). We anticipate simple time trends (e.g., linear or quadratic) for the mixed effects models; however, we will use linear spline models in the event complex time trends are observed, such as non-linear trends that cannot be approximated by simple polynomials in time [2]**.** These analyses will be performed using PROC GLIMMIX.

Primary Hypothesis 2: Clients in the PATH condition will demonstrate greater retention in treatment than clients in the SC condition. Retention is widely considered the leading proxy for clinical effectiveness in SUD treatment. A Cox proportional hazards survival analysis [5,6] will be used to determine whether the rate of dropout differs between the two treatment conditions. The analysis will account for right-censored data as some participants may not experience the event (i.e., treatment dropout, defined as absence from treatment for 30 days) during the period of observation. The analysis will include any necessary covariates. To establish the HTE for this outcome, the survival analysis will be repeated for each primary and exploratory subgroup identified above. Models will include a term for the subgroup variable and its interaction with treatment condition. Again, if the interaction term is significant, the treatment effect will be estimated separately at each level of the categorical variable that is used to define the subgroups. We anticipate that the proportional hazard assumption will be met as we expect the typical pattern of dropout in both groups (i.e., high hazard early, low hazard late). We will conduct standard residual-based analyses to evaluate proportionality and, in the event that the

proportional hazard assumption is violated, we will use extended Cox models that include a time-dependent covariate to model the non-proportionality [7]. The analyses will be performed using PROC PHREG.

Power for the Primary Hypotheses. Power analyses were calculated for the primary hypotheses described above. The power analyses for Primary Hypothesis 1 were calculated following the recommendations of Liu and Wu [8]. The analysis was based on a corrected alpha of .025 (.05/2 primary outcomes), a compound symmetry covariance structure, an estimated abstinence rate of 25% in the SC condition (based on data from our prior study), a correlation between observations on the same patient of .50, and assuming 20% attrition at the final time point. For the smallest subgroup (patients primarily using prescription opioids: n = 266; 133 per condition), we will have 80% power to detect an approximately 15% difference in abstinence rates between the conditions (i.e., 25% in control vs. 40% in experimental; OR = 2.0). This projected difference is conservative and based only on the repeatedly demonstrated, similarly scaled differences contingency management yields when compared to standard treatment conditions [9-13]. The smallest detectable differences for the remaining subgroups and the overall sample are as follows using the same specifications: (1) patients recruited from IP setting (n = 400): 12%, OR = 1.8, (2) patients recruited from OP setting (n = 400): 12%, OR = 1.8, (3) patients primarily using heroin (n = 534): 9%, OR = 1.6, and (4) the overall sample (n = 800): 8%, OR = 1.5. The power analyses for Primary Hypothesis 2 were calculated following the recommendations of Lakatos [14]. The analysis was based on a corrected alpha of .025, no attrition as chart data will be available for all patients, and estimated dropout rate of 60% in the SC condition [15]. For the smallest subgroup (patients primarily using prescription opioids: n = 266; 133 per condition), we will have 80% power to detect a hazard ratio of .59 which corresponds to a difference in survival rates of approximately 18% between the two conditions. The smallest detectable hazard ratio for the remaining subgroups and overall sample are as follows using the same specifications: (1) patients recruited from IP setting (n = 400): HR = .65, 15% difference in survival (2) patients recruited from OP setting (n = 400): HR = .65 , 15% difference in survival (3) patients primarily using heroin (n = 534): HR = .69, 13% difference in survival , and (4) the overall sample (n = 800): HR = .74, 10% difference in survival. We have chosen to use an alpha level that corrects for the number of primary hypotheses rather than the number of subgroup analyses as the latter approach would substantially increase the likelihood of making a type 2 error [4,16].

Secondary Hypotheses: Relative to those in the SC condition, participants in the PATH condition will demonstrate:

- ***Lower rates of service utilization (days of hospitalizations, emergency department visits, residential treatment)***
- ***Higher quality of life (Q-LES-Q-SF scores)***
- ***Lower rates of HIV risk behaviors (RAB total, sex, and drug risk scores)***

Exploratory Hypotheses: Relative to those in the SC condition, participants in the PATH condition will demonstrate:

- ***Improved Employment, Family/Social Functioning, and Psychiatric severity scores on the ASAM Criteria.***

Mixed effects models [2] will be used to compare the treatment conditions on each secondary and exploratory outcome at months 6, 12 and 18. Each model will include terms for treatment condition, time, and their interaction, and the baseline value of outcome measure along with any other necessary covariates. Subgroup analyses similar to those outlined above will be conducted to establish the heterogeneity of the treatment effects. Count data (i.e., service utilization variables) will be analyzed using PROC GLIMMIX specifying a poisson distribution and continuous data (i.e., quality of life, HIV risk scores, and psychosocial severity scores) will be analyzed using PROC MIXED.

Missing values. Missing data will not be an issue for some of the outcomes (i.e., time in treatment). The most important source of missing data in this study is likely to be from participant dropout. We will analyze the data under an intent-to-treat principle in which all participants initially randomized will be included [17]. In addition, mixed effects models will provide valid estimates of intervention efficacy if the missing data are ignorable. We will evaluate the patterns of missing data for each outcome and conduct logistic regression and survival analysis models predicting dropout using client characteristics and accrued measures to determine whether the dropout process is ignorable (i.e. whether missed visits are well explained by observable data). We will examine the impact of missing data by comparing results from models in which missing data are imputed and through pattern-mixture models [18] that examine how responses vary as a function of the pattern of missing data.

Sensitivity analyses . We will conduct a series of sensitivity analyses for the primary hypotheses as recommended by Thabane et al. [19]. We will examine the influence of outliers (identified through boxplots and z-scores) by comparing results from analyses with and without the outlying values. We will examine the impact of protocol non-compliance by comparing the results from the intent-to-treat analyses to analyses in which only participants who received a sufficient dose or different components of the intervention are included; protocol adherence will be represented by weekly medication adherence scores, weekly counseling scores, and combined scores for the sensitivity analyses calculated from TSR and chart review. Furthermore, we will extend the models described above for our primary and secondary outcomes using instrumental variable approaches. Here, the instrumental variable is treatment condition and, under some conditions, will control for unmeasured bias (due to participants self-selecting to comply or not) when estimating the relative effects of the interventions in those who adhere to their assigned intervention. Thus, the results will use data from all subjects who show various degrees of compliance and will yield estimates of the intervention effects that would have been seen in a study with full compliance. For a single overall measure of compliance, the methods of Nagelkerke et al. [20] will be used; if there is sufficient within-subject variation across time in compliance, the longitudinal methods of Small et al. [21] will be used. These sensitivity analyses allow us to determine whether estimates of effect from the primary models can be reported without explicit adjustments for protocol non-compliance. We will examine the impact of our distributional assumptions by comparing results obtained under different distributional assumptions (e.g., parametric vs. non-parametric). We will use multiple imputation and pattern-mixture models to evaluate the impact of missing data as discussed in the prior section. Finally, we will examine the impact of using different definitions of outcomes (e.g., different cutoff points for binary outcomes, objectively verified *vs.* self-reported substance use) by comparing results obtained from models using different operationalizations of the

outcome variables. These sensitivity analyses will allow us to evaluation the robustness of the study results.

***Heterogeneity of treatment effects*.** The subgroup analyses described above will be used as a test of the heterogeneity of treatment effects. In addition, we have identified three patient characteristics that may be related to treatment effects: psychiatric problem severity (meets pre-determined threshold vs. does not), OUD problem severity (meets pre-determined threshold vs. does not), and housing status (homeless or housed). Each of these variables should negatively impact treatment outcomes, and could suppress differences between conditions. To evaluate the heterogeneity of treatment effects as a function of psychiatric severity, the mixed effects models and survival analyses used to test the primary hypotheses will be modified to include a binary term reflecting psychiatric severity and the interaction of psychiatric severity and condition. A significant interaction would indicate that the effects of the intervention vary as a function of a patient's psychiatric severity (i.e., treatment effects are heterogeneous). If the interaction term is significant, the treatment effect will be estimated separately at each level of the categorical variable that is used to define the subgroups. Similar analyses will be performed for OUD problem severity and housing.

References

1. Hauck WW, Anderson S, Marcus SM. Should we adjust for covariates in nonlinear regression analyses of randomized trials? *Controlled Clinical Trials.* 1998;19(3):249-256.
2. Fitzmaurice GM, Laird NM, Ware JH. *Applied Longitudinal Analysis (2nd Edition).* Wiley Series in Probability and Statistics; 2011.
3. Diggle PJ, Heagerty PJ, Laing KY, Zeger SL. *Analysis of longitudinal data (2nd Ed.).* Oxford, England: Oxford University Press; 2002.
4. Varadhan R, Seeger JD. Estimation and reporting of heterogeneity of treatment effects. In: Velentgas P, Dreyer NA, Nourjah P, Smith SR, Torchia MM, eds. *Developing a protocol for observational comparative effectiveness research: A user's guide*: Government Printing Office; 2013.
5. Hosmer DW, Lemeshow S, May S. *Applied survival analysis. Second edition.* Hoboken, NJ: John Wiley and Sons, Inc.; 2008.
6. Klein JP, Moeschberger ML. *Survival analysis: Techniques for censored and truncated data. Second edition.* New York, NY: Springer Science & Business Media; 2005.
7. Allison PD. *Survival analysis using SAS: A practical guide.* SAS Institute; 2010.
8. Liu H, Wu T. Sample size calculation and power analysis of time-averaged difference. Journal of Modern Applied Statistical Methods. 2005;4(2):434-445.
9. Rawson RA, Huber A, McCann M, et al. A comparison of contingency management and cognitive-behavioral approaches during methadone maintenance treatment for cocaine dependence. *Archives of General Psychiatry.* 2002;59(9):817-824.
10. Griffith JD, Rowan-Szal GA, Roark RR, Simpson DD. Contingency management in outpatient methadone treatment: A meta-analysis. *Drug Alcohol Depend.* 2000;58(1-2):55-66.
11. Lussier JP, Heil SH, Mongeon JA, Badger GJ, Higgins ST. A meta-analysis of voucher-based reinforcement therapy for substance use disorders. *Addiction.* 2006;101(2):192-203.
12. Prendergast M, Podus D, Finney J, Greenwell L, Roll J. Contingency management for treatment of substance use disorders: A meta-analysis. *Addiction.* 2006;101(11):1546-1560.
13. McKay JR, Lynch KG, Coviello D, et al. Randomized trial of continuing care enhancements for

cocaine-dependent patients following initial engagement. *J Consult Clin Psychol.* 2010;78(1):111-120.
14. Lakatos E. Sample sizes based on the log-rank statistic in complex clinical trials. *Biometrics.* 1988;44:229-241.
15. Substance Abuse and Mental Health Services Administration Office of Applied Studies. Treatment Episode Data Set (TEDS): 2006. Discharges from substance abuse treatment services. Rockville, MD DASIS Series: S-46, DHHS Publication No. (SMA) 09-4378; 2009.
16. Drachman D. Adjusting for multiple comparisons. *J Clin Res Best Pract.* 2012;8:1-3.
17. Lavori PW. Clinical trials in psychiatry: Should protocol deviation censor patient data? *Neuropsychopharmacology.* 1992;6(1):39-48; discussion 49-63.
18. Little RJA. Modeling the drop-out mechanism in repeated measures studies. *Journal of the American Statistical Association.* 1995;90:1112-1121.
19. Thabane L, Mbuagbaw L, Zhang S, et al. A tutorial on sensitivity analyses in clinical trials: The what, why, when and how. *British Medical Research Methodology.* 2013;13:92.
20. Nagelkerke N, Fidler V, Bernsen R, Borgdorff M. Estimating treatment effects in randomized clinical trials in the presence of non-compliance. *Stat Med.* 2000;19(14):1849-1864.
21. Small DS, Ten Have TR, Joffe MM, Cheng J. Random effects logistic models for analysing efficacy of a longitudinal randomized treatment with non-adherence. *Stat Med.* 2006;25(12):1981-2007.