Reducing Offenders' HIV Risk: MI Enhanced Case Management With Drug-Free Housing
Data Plan, Version October 7, 2016
Clinicialtrials.gov ID: NCT01977092

Analysis

Univariate and bivariate analyses of all variables will inform development of longitudinal models to follow. These analyses will include identification of outliers and data distributions that indicate need for data transformation procedures or categorization of variables. These procedures will be important to avoid use of covariate variables that would trigger common violations of modeling assumptions. Bivariate analyses will include histograms, scatterplots, correlation matrices, χ^2 tables and ANCOVA models for all study variables. Multiple testing procedures (e.g., Bonferroni) will be used, when appropriate, to prevent inflation of type I error rates.

Longitudinal Analysis of Outcome Measures

Our key analytic tools for comparison of study conditions across time will be latent profile analysis (Diggle, et al., 2002) and latent growth curve modeling (Raudenbush & Bryk, 2002). Latent profile analysis will first be used to inform the appropriate choice for parameterization of the growth curve model as it simply estimates the mean outcome values at each interview while also incorporating non-independent variances structures within individual over time. Should the mean outcomes across time appear to follow a parametric response function (e.g., linear or linear after transformation, or quadratic), growth curve modeling will then be implemented. Such a growth curve model provides the flexibility of allowing each individual to have their own growth curve parameters (e.g., intercept and slope) and thus longitudinal trajectory for each outcome variable analyzed (e.g., days of substance use, arrests, and incarceration over the treatment period). As respondents will be randomized, baseline differences are not expected on outcome variables. Should any exist however, time-invariant covariates used as predictors of the random intercept should be able to identify them. A key capability of growth curve modeling is its ability to test for the significance of variability in outcome measures between individuals as well as treatment and control groups. Time-invariant control variables will be included in the formulae for both the random intercept and slope. Examples include gender, age, marital status, ethnicity, and any other key baseline variables that may influence the outcome across time. SPSS and MPlus will be used to carry out these analyses (Muthén & Muthén, 2011; SPSS Inc., 2009).

Hierarchical Linear Modeling (HLM)

We will use HLM to assess whether the 32 houses studied differ on primary outcome measure (e.g., HIV risk, alcohol and drug use on the TLFB, and arrests) as well as overall scores on the SMPH. An example of a growth curve model that will be estimated, after first inspecting results from longitudinal profile analyses to inform the choice of the most appropriate choice for growth curve parameterization, will be a 3-level model with time nested within respondents nested within houses, a simple example of which will be $y_{t,i,h} = \alpha_{i,h} + \beta_{i,h}t + \varepsilon_{t,i,h}$ where the random intercept for the i^{th} individual in house h will be parameterized as $\alpha_{i,h} = \vartheta_{0,h} + \vartheta_1 X_{i,h} + u_{0,i,h}$ and the random slope will be parameterized as $\beta_{i,h} = \lambda_{0,h} + \lambda_1 X_{i,h} + u_{1,i,h}$ where $X_{i,h}$ are time invariant individual-level covariates (e.g., baseline age, ethnicity) and the $u_{o,h}$ and $u_{1,h}$ are random variables that indicate unexplained variation in the random intercept and slope across individuals. To assess variability of the dependent variable across houses, both at baseline (the intercept) and over time (the slope), the random intercept of the level 1 intercept ($\theta_{0,h}$) and the random intercept of the level 1 slope ($\lambda_{o,h}$) will be parameterized as $\vartheta_{0,h} = \gamma_{0,0} + \gamma_{0,0}Z_h + v_{0.h}$ and $\lambda_{0,h} = \gamma_{1,0} + \gamma_{1,0}Z_h + v_{1.h}$ where the Z_h are house-level variables and the $v_{o,h}$ and $v_{1,h}$ are random variables indicating unexplained variation in the average growth parameters across houses. Therefore, both individual-level and house-level variables will be incorporated into the longitudinal model to predict variability in growth parameters. In

addition to accounting for the nesting of individuals within time, the models will account for the nesting of individuals within houses which, when left unconsidered even if no house level variables *Z* are entered into the model, may result in underestimates of level 2 model variances u_o and u_1 and therefore invalid inferences for the effects of individual-level variables on longitudinal trajectories.

Dose-Response Analysis for MMI

For participants in the MMI condition we will conduct a dose-response analysis. Our rationale draws on the case management study by Longshore et al (2005). They reported a significant relationship between the number of case management contacts with residents and primary outcome variables. We will complete a similar analysis using bivariate associations (e.g., by examining χ^2 tests and estimated correlations) and multivariate models that assess dose effects controlling for selected variables.

Analysis of Specific Aims

Aim 1: *To assess HIV testing rates, HIV risk behaviors, and utilization for HIV services at baseline, 6, and 12 months.*

H1.1 Residents receiving MMI in addition to drug free housing will have higher rates of HIV testing, higher utilization of HIV services, and fewer HIV risk behaviors than the drug free housing and resource manual condition.

Hypotheses H1.1 will be tested in the framework of longitudinal models, as described above. The focus is on treatment condition which, although uniquely assigned to individuals, is implemented at the house level and therefore will be treated as a level 3 variable. This allows for a limited exploratory examination of potential treatment effects with house level characteristics. Interactions of treatment condition with level 2 variables (e.g., gender) can be implemented by simply including treatment as a predictor of the non-randomly varying level 2 coefficient (e.g., gender) of interest. Evidence in favor of the hypotheses will be significant level-3 coefficients (the relevant γ coefficient) for the indicator variables (Z) representing average differences between houses in the rate of change in the outcome over time (i.e., the intercept of the level 2 slope parameter $\lambda_{o,h}$).

Aim 2: *To compare baseline substance use within each study condition with 6- and 12-month substance use.*

H2.1 Residents in each study condition will show significant reductions in drug and alcohol use *between <u>baseline and follow-up time points</u>* as measured by the Time Line Follow Back (TLFB).

H2.2 Residents in each study condition will show significant reductions in drug and alcohol problems *<u>between baseline and follow-up time point</u>s* as measured by the Addiction Severity Index alcohol and drug scales.

Hypotheses H2.1 and H2.2 will be tested in the framework of longitudinal models, as described above. As a first step, an empty model will be estimated (random intercept and slope but no level 2 (individual) or level 3 (house) covariates) to test whether the variances of the intercept and slope parameters (*u* and *v)* are significant. Should significant variation be found, candidate covariates, both at the individual level (level 2) and the house level (level 3), will be considered in separate models and any such covariates will be entered as predictors of both the random intercept and slope.

Aim 3: To compare outcomes between the two study conditions at baseline, 6, and 12 months.

H3.1 Residents receiving MMI in addition to drug free housing will have less substance use (TLFB) and lower ASI alcohol and drug severity than the drug free housing and resources condition.

H3.2 Residents receiving MMI in addition to drug free housing will have fewer arrests, fewer days incarcerated and lower ASI legal severity than the drug free housing and resources condition.

H3.3 Residents receiving MMI in addition to drug free housing will have fewer work problems (ASI Employment scale) and more days worked than the drug free housing and resources condition.

Testing of hypotheses in H3.1 to H3.3 will first be carried out using the same growth curve modeling methods as described for analyses proposed in Aim 1 above.

Aim 4: To assess mediators and moderators of the MMI.

H4.1 Service utilization (Hser et al., 1991) will mediate the relationship between treatment type (drug free housing with vs. without MMI) and outcome (TLFB, ASI Alcohol and Drug scales and criminal justice recidivism).

The best fitting model found from analyses in Aim 2 which include level 3 indicators of treatment condition as predictors of the average random slope (and potentially the random intercept) will serve as the baseline model with which to assess mediation. For treatment predictors of the random slope, overall significance of the 1 degree of freedom test of differences in changes over time between the 32 houses will be assessed using a linear contrast test. Then, time varying covariates (e.g., utilization of services, perceived stress, and negative affect scales) will be entered separately into the model as a level 1 (time within individual) predictor. Then, the same linear contrast test of the level 3 treatment coefficients for the random slope coefficients indicating average differences in change over time between treatment conditions will again be carried out, with reductions in the magnitude of the coefficient estimate and significance of the test indicating the degree to which such variables mediate program differences.

Exploratory Aims:

We will explore a variety of additional outcomes within and between study conditions for which we do not have a priori hypotheses. These will include additional ASI scales (Family, Medical, and Psychiatric) and Housing Status. We will also explore how a variety of covariates are associated with outcome within and across study conditions. Variables assessed will include social support, supportive confrontation, motivation, drug use, age of onset for alcohol and drug use, criminal justice history and status (probation or parole) and gang affiliation. Analytic methods will include first examining univariate distributions and bivariate associations and then entering variables as covariates into the growth curve modeling described above.

A second exploratory aim is to conduct 30 qualitative interviews (15 in each condition) to identify general and intervention specific factors that were perceived to affect outcomes and ways to improve the two active interventions. An initial coding list will include specific codes suggested by characteristics of the intervention. For example, codes for the SLH alone condition might include things like “helped me increase support to not use alcohol or drugs,” and “helped me avoid exposure to alcohol and drugs.” For MMI, codes might include “made suggestions for services” and “helped me prepare for living at the SLH.” For meditation, codes might include things like “helped me relax” and “increased positive feelings about myself.” Using line-by-line coding, themes and excerpts of text will be highlighted as quotations that exemplify salient resident attitudes and beliefs. Excerpts will be printed out, and each analyst will group the excerpts into thematic categories. The extent of agreement between the analysts’ initial coding of the data will be assessed. Each analyst will define the working themes and through discussion select a final set. Data will be entered into the qualitative software program, N Vivo. The research team will use the constant comparison method and analytic memoing to identify patterns, create themes or categories, and develop theoretical insights (Strauss & Corbin, 1990).

Power Considerations

Given our sample of 165 in each of the 2 study conditions and our expectation that we will contact 80% of the sample at 12 months we should have 132 per group and a total N of 264 available for analysis for at least two time points. For tests of equality of paired differences in means for outcomes between the final interview and baseline between the two study conditions, power to detect small-to-medium effect sizes (d=.35) would be .82 in 2-sided tests with α=.05.

Similar power is available for tests of equality of proportions (Cohen, 1988).

For estimation of power in growth curve modeling, programs for estimating power are available (Bosker, 1999), but they typically require population parameters, including between- and within-subjects covariance matrices and estimates of level 1 and 2 residual variances. Other methods have been developed for testing specific slope differences (Jung & Ahn, 2005), but point estimates are still required for the variances of the slope parameter within each group for which prior information, and even ranges of plausible values, are simply not available. Instead of assuming knowledge of such population parameters, power is approximated here in the context of longitudinal Generalized Estimating Equations (GEE) models (Diggle, et al., 2002) using the attrition adjusted effective sample sizes. Various possible error variance structures in GEE models can approximate the partitioning of variance explicit in multilevel models and also provide a simpler context in which to approximate power for between groups longitudinal effects as it instead relies only on an effect size defined in standardized units as $\Delta = d/\sigma$, where d effect size (here, the time averaged difference in the outcome variable between the two comparison groups) and σ is the unknown level 1 error variance. Assuming α=.05 and n=3 interviews per respondent, the minimum sample sizes required to detect standardized effect size differences of Δ=(.2, .3, .4, .5) for exchangeable error correlation values of δ=(.2, .5, and .8) with power of .80 are shown in the table above. Power for estimation of individual level effects associated with the first two levels comprising interviews nested within individuals uses a design effect adjusted sample size accounting for within-house correlation of responses. Assuming a medium Intraclass Correlation Coefficient of .06 (Kirk, 1995) reduces the effective attrition adjusted number of respondents from 132 to n=132/(1+32*.06)=45 cases. From the table, power would be close to .80 to detect small-to-medium effect size differences (Δ=.35, as defined in (Cohen, 1988) in the time averaged differences between the two treatment groups when the error correlation over time is relatively moderate (δ =.2). Power for estimation of level 3 coefficients is approximated in the context of simple linear regression with n=32 cases (houses) available. Power to detect medium to large effect size relationships (f^2=.25) (Cohen, 1988) between house level characteristics and time averaged differences in outcomes between the two treatment conditions would be .81.

Min N for Power=.8 in Longitudinal Models with Continuous Outcomes

		Δ (%)			
		.2	.3	.4	.5
δ	.2	146	65	37	24
	.5	208	93	52	34
	.8	270	120	68	44

The meta-analysis on MI outcome by Hettema et al. (2005) found pairwise between group effect sizes to be .77 (very large) in the short term and .30 at (small-to-medium) one year follow up. Therefore, power should be more than adequate.

Management Plan

The management plan is based on procedures from our recent study of SLHs (i.e., Polcin et al., 2010) as well as previous multisite studies where Dr. Polcin served as key personnel (Carroll et al., 2006). Coordination will be achieved using weekly telephone conference calls involving all staff on the research team. Dr. Polcin and Ms. Evans will travel to Los Angeles at least quarterly for day long face to face meetings with fieldwork staff and therapists. Study therapists will be hired by Dr. Polcin who will train them together (16 hours face to face) and meet with them via telephone conferencing weekly to discuss all aspects of MICM.

Timeline

Months 1-4	Recruit and train therapists and interviewers, IRB approval, Federal Certificate of Confidentiality, Review instruments and data collection procedures, Complete Procedures Manual and Question-by-Question guidelines, Pilot test interventions and data collection procedures.
Months 5-43	Recruit approximately 8 study participants per month over 38 months. Finish data cleaning for baseline data at month 43.
Months 44-52	Complete all follow up interviews.
Months 53-60	Complete longitudinal analysis of data. Submit papers to addiction and criminal justice journals as well as trade magazines for addiction professionals. Present final analyses to consumers and other stakeholders.