Statistical analysis plan for ACF/OPRE grant #90FA3010 entitled, "KEEP Connecting Kin: Improving Outcomes for Informal Kinship Care Families Via an In-Home Parenting and Peer-Support Program"

Document created on 7/11/2023

NCT # (yet to be assigned)

<div align="center">**STATISICAL ANALYSIS PLAN**</div>

Adherence to the Clearinghouse Standards for High Support of Causal Inference.

The KEEP-CK study utilizes a rigorous waitlist-controlled "services as usual" (SAU) randomized design, and implementation, assessment, measurement, and analytic methods that adhere to the Prevention Services Clearinghouse's standards for high support of causal evidence.

Integrity of random assignment and blocking. Informal kinship caregivers will be randomly assigned to receive the KEEP-CK program either immediately or after waiting 10 months using a randomized blocked design. The integrity of random assignments will be maintained throughout the study period (e.g., participants who dropout of the program will be retained in the analyses, randomization will occur for all participants across all six agency providers, the KEEP-CK groups will not be "filled" with non-randomized study participants). In addition, to minimize the possibility of both differential attrition across the intervention and waitlist control groups, and potential differences between the two groups at baseline, <u>three randomization blocking factors will be utilized</u>, including: (a) the length of time that the child has been living in informal kinship care (6 months or less, 7 months to less than 2 years, or 2 or more years); (b) child age (4 to 12 years or 13+ years); and (c) low (or not low) resource families, defined by an annual household income below (or above) 150% of the federal poverty level. These blocking factors were selected based on the Prevention Services Clearinghouse's requirement of evaluating baseline equivalence on socioeconomic status and child age, and to account for potential differences in kinship caregivers' urgency for receiving supports and services due to having recently started informally caring for a child.

High causal support for impact analyses and measures. The impact analyses of the KEEP-CK program will also adhere to high standards for causal inference. Although not expected, if an unequal number of participants are allocated to the intervention and waitlist control conditions within randomization blocks, this will be accounted for in the impact analyses. In addition, if baseline equivalence cannot be established across the two groups, an acceptable method for controlling for

baseline differences will be employed in the impact analyses (e.g., including baseline variables as covariates, repeated measures analysis of variance). Finally, endogenous covariates (i.e., covariates measured after baseline and that could have been affected by the intervention) will not be included in the impact analyses. Further, all measures for the current study will adhere to the Clearinghouse's measurement standards for high causal evidence. This includes having face validity, high reliability (Cronbach's alpha of at least 0.5), and consistency in how and when measures are collected and constructed across the intervention and SAU waitlist control groups.

Minimizing design confounds. Design confounding factors that can be threats to causal inference will be minimized by using a randomized blocked waitlist control design. Further, given that six agencies across the state of Oregon will provide KEEP-CK services to informal kinship families, we do not anticipate that there will be a "*substantially different characteristics confound*" across the intervention and control groups due to providers—nor will there be a "*n=1 administrative unit provider confound*" given our ongoing partnerships and relationships with the six agency providers. Likewise, at least 24 KEEP-CK group leaders will be trained and certified to lead the groups and deliver the intervention, thus the current study design is not subject to a "*n=1 person-provider confound*".

Minimizing and accounting for missing data. The study team is highly experienced and skilled in recruitment and retention efforts, and strategies will be employed to minimize missing data (e.g., personalized contact, adequate compensation for participants' time). However, youth living in informal kinship care are highly mobile and there will be some missing data. Full information maximum likelihood estimation[1] will be used for parameter estimation, which is known to yield unbiased parameter estimates provided the data are missing at random (MAR[2-4]). Multiple imputation may also be used for variables with high amounts of missingness or independent variables.[4] Missing data correlates may also be incorporated in Mplus to make the MAR assumption more attainable.[5]

In sum, this study adheres to all of the Prevention Services Clearinghouse's requirements for high support of causal evidence, including randomization to the intervention and SAU waitlist control

conditions with standards for integrity of random assignment, attrition, baseline equivalence, statistical modeling, measurement, design confounds, and missing data.

Analytic Strategy.

Any necessary adjustments described above for high causal support will be incorporated into the hypothesis tests. The general analytic procedure will be to calculate slopes (changes in the outcomes from pre- vs. post-intervention and pre-intervention vs. follow-up) for the targeted intervention outcomes of child well-being, parenting behaviors and parent/caregiver stress (Aim 1), and service utilization outcomes (Aim 2). To test our research hypotheses, the parent- or youth-level slopes then become the outcomes of interest and are, for example, modeled as a function of KEEP-CK versus waitlist control SAU group status to determine the impact of the program. For child permanency (Aim 1) and entry into the CWS (Aim 3), we will examine differences between the two groups at the post-intervention and 10-month follow-up assessments. Analytical models will account for non-independence of observations, missing data (described above), attrition, and noncompliance.

Clustering and nested data. The proposed study design is longitudinal and clustered. Repeated assessments across time (level-1) are nested within kinship parents/youth (level-2), half of which are nested within KEEP groups (or received SAU) (level-3). In all analyses, dependence among scores due to clustering effects will be accounted for by adjusting the standard errors using sandwich estimator[5] or using multilevel modeling (MLM).[6] Thus, we will account for the nested structure of the data.

Randomization. This will occur at the level of the kinship parent–youth dyad using a randomized blocked waitlist control design (see above for the blocking factors).

Validity and reliability of the measures. We will first evaluate the overall reliability and validity of the measures by conducting exploratory and confirmatory factor analyses. Next, we will test for measurement invariance[7] across a number of factors (e.g., intervention status, sex, assessment time points).

Analyses by aim. All models will be estimated using MLM to account for the multilevel structure of the data and estimated in Mplus.[5,6] Both maximum likelihood estimation and bootstrapped 95% confidence intervals will be estimated. The Type I error rate will be adjusted for multiple testing. Potential differences in the targeted child outcomes attributable to child age, biological sex, and race/ethnicity will be controlled for while testing the research questions, which evaluate the impacts of the KEEP-CK intervention on child and adult outcomes (Aim 1), service utilization outcomes (Aim 2), and child entry into the CWS (Aim 3).

Impact analyses. For child well-being, parenting practices and parent/caregiver stress (Aim 1), and service utilization outcomes (Aim 2), the immediate and long-term impacts of the KEEP-CK intervention will be evaluated by estimating the group status (KEEP-CK vs. SAU waitlist control) by time (pre- vs. post-intervention or pre-intervention vs. follow-up) interaction. Thus, the posited benefits of providing services to kinship families will be evaluated by examining whether outcomes are significantly improved for those who received KEEP-CK services compared to those randomly assigned to the SAU waitlist control condition—and further, whether any such benefits are sustained 6 months after the end of the intervention. Youth internalizing and externalizing outcomes will be evaluated in terms of overall symptoms (i.e., CBCL and YSR broadband *t*-scores) and borderline and clinical cutoff scores (i.e., binary indicators). Service utilization outcomes will be evaluated in terms of the overall number of services utilized and each specific type of service.

To evaluate the impact of the KEEP-CK on child permanency (Aim 1), mean differences between the intervention and SAU waitlist control groups will be compared at the post-intervention and follow-up assessment for placement stability, and at the 10-month follow-up assessment for permanency of placements. Likewise, to evaluate whether KEEP-CK prevents entry into the CWS (Aim 3), we will use logistic regression to examine if youth are less likely to enter into child welfare custody at the 10-month follow-up assessment if their kinship parents had participated in KEEP-CK compared to youth whose caregivers had yet to receive services.

Qualitative analyses of parent and youth perspectives on KEEP-CK (Aim 4).

Clark (Co-I), an expert coder, will code all of the focus group transcripts using ATLAS.ti software version 23. A codebook will be developed using the primary questions from the focus group or interview protocol for broad themes and identifying emergent subthemes over the course of transcript analysis. Coding will involve the initial assignment of portions of text to one or more codes in the codebook, updating the codebook throughout the process, and comparing all transcripts against the final codebook.[8] To ensure scientific rigor and credibility of findings, a second expert coder (Buchanan, Co-I) will review the coded work. We will conduct member checks with participants as needed to ensure accuracy of data and interpretation. Both coders will participate in a peer debriefing process to identify representative quotes for the study results.

REFERENCES

1. Arbuckle JL. Full information estimation in the presence of incomplete data. In: Marcoulides GA, Schumaker RE, eds. *Advanced structural equation modeling: Issues and techniques*. Lawrence Earlbaum Associates; 1996:243-277.

2. Little RJA, Rubin DB. The analysis of social science data with missing values. *Sociological Methods and Research*. 1989;18:292-326.

3. Muthén BO, Kaplan D, Hollis M. On structural equation modeling with data that are not missing completely at random. *Psychometrika*. 1987;52(3):431-462.

4. Schafer JL, Graham JW. Missing data: Our view of the state of the art. *Psychological Methods*. 2002;7(2):147-177. doi:10.1037/1082-989X.7.2.147

5. Muthén LK, Muthén BO. *Mplus user's guide*. Muthén & Muthén; 1998–2017.

6. Raudenbush SW, Bryk AS. *Hierarchical linear models: Applications and data analysis methods*. 2nd ed. vol 1. Sage; 2001.

7. Meredith W. Measurement invariance, factor analysis, and factorial invariance. *Psychometrika*. 1993/12/01 1993;58(4):525-543. doi:10.1007/BF02294825

8. Padgett DK. *Qualitative methods in social work research*. Sage; 2008.