All in Dads! Healthy Marriage and Responsible Fatherhood Program Evaluation

Statistical Analysis Plan

NCT05194020

October 17, 2024

1. Data Analysis

Briefly describe the planned approach for data analysis. If an impact analysis is proposed, name the key dependent and independent variables, and describe any methods to minimize Type I error (i.e., finding positive impacts by chance) such as limiting the number of impacts to be analyzed and/or multiple comparison correction. Describe proposed approach(es) for addressing missing data.

Identify study sample

The study sample includes participants who: 1) gave recorded informed consent to be part of the study, and 2) took both an entrance and exit survey (for attitudinal outcomes) or entrance and follow-up survey (for behavioral outcomes). The CONSORT diagram in Figure 2 shows exclusion criteria and the anticipated size of the overall study sample. The analytic sample for particular outcome measures will vary based on missing data.

Defining analysis measures

For each outcome construct, we have identified the relevant measures from nFORM that are theoretically aligned with that construct. We will generate a correlation matrix between items in a given construct to ensure that theoretically related items are also empirically related in our data set. Items that are not strongly correlated with other items in a construct will be removed as necessary. Factor analysis will be used to ensure that all construct items hang together (using an alpha of 0.7 or higher as the threshold). If an alpha of 0.7 cannot be obtained, that outcome will be removed.

Once we are confident that all of the items align with a given construct, we will create a composite measure by taking an average of the scores on each non-missing item in the construct. The measure definition matrix below provides information on each proposed outcome. For composite measures, a change score will be calculated between a participant's pre-program composite score and post-program composite score. For standalone survey items, a change score will be calculated between a participant's pre-survey and post-survey responses.

For the sake of consistency and simplicity, parenting and co-parenting analyses will use a participant's youngest child as the focal child.

Model Specification to Predict Outcomes With Dosage

Construct by Outcome	Dependent Variables (follow-up)	Response Codes for Dependent Variables	Dependent Variable Type	Independent Variables
Primary outcome:				
Parenting Behavior: 1 item	HOURSFU	# hours spent with child	Continuous	Dosage: hours of attendance per service
	REACHFU	1 = Almost every day to 5 = never in past month		

				(using nFORM data); Co-variates: development starting points, risk - family structure, demographics, socioeconomics (ACS & OLLE survey data at baseline)
1 item 11 scaled items	ENGAGFU	1 = never to 5 = every day or almost every day		
Co-parenting Behavior: 11 scaled items	COPARFU	1 = strongly disagree to 5 = strongly agree		
		1 = yes, 0 = no		

Secondary outcome:

				Dosage: hours of attendance per service (using nFORM data); Co-variates: development starting points, risk - family structure, demographics, socioeconomics (ACS & OLLE survey data at baseline)
Parenting Attitudes: 7 scaled items	PATTFU	1 = always to 5 = never	Continuous	
	SAVFU			
Financial Readiness 1 item 1 item	CHECKFU	1 = yes, 0 = no	Dichotomous	

Handling missing data

Outcomes

When creating the composite measures for co-parenting and parenting relationships, we will create a composite score by taking the average of multiple individual items. For these measures, our current plan is to use 20% as a threshold for allowable missing items, based on guidance from evaluation technical assistance resources. This plan is contingent on the final distribution of missing data in our data set. With a 20% threshold, if a respondent is missing more than 1 item used for creating the co-parenting construct or more than 2 items for creating the parenting construct, the respondent will be assigned a missing value for that construct. We will not be imputing truly missing values for outcomes. To create a construct score, the average will divide by the number of non-missing values in the construct.

Participants who have not seen their child within the past month will have their parenting behavior responses set to "Never" (1) and will be included in the analytic sample for that outcome.

For outcomes that use single survey items, participants who do not respond to the item on either the pre- or post-survey will be excluded from the analytic sample for that outcome.

Missing data for the implementation outcomes could happen as a result of programmatic data entry issues. For the sake of this evaluation, we assume that any primary or support services received by a participant are being accurately logged into nFORM, so a lack of documented attendance, service contacts, referrals, etc. is indicative of lack of services (i.e., a participant not receiving the program components as intended).

Assessing non-response bias

We will conduct response rate analysis for each primary outcome of interest to assess non-response bias and adjust for threats to internal validity. Using data from the Applicant Characteristics Survey, we will look at demographics (race, ethnicity, age, education level) and primary reason for joining the program among participants who fall into each of the following categories: 1) non-respondents who answered no surveys after the ACS, 2) respondents who completed the pre-survey only, 3) respondents who completed the post-survey only, and 4) respondents who completed both the pre- and post-survey (complete case).

Analytic approach

The main goal of this descriptive study is to assess pre-post change scores in the attitudinal outcomes listed above and assess pre-to-one-year follow-up change scores in the behavioral outcomes listed above among program participants before and after primary workshops. We will use paired sample t-tests to assess the magnitude and significance of changes among program participants in the analytic sample for each outcome. For binary economic stability outcomes, we will descriptively report the percent of participants who opened a checking or savings account one year after enrollment, compared to the percent who had opened these accounts before the program. We will adjust our p-values for multiple hypothesis testing and report the adjusted p-values in the appendix of the final report.

2. Data Archiving and Transfer

Briefly describe the planned approach data arching and transfer by addressing questions below.

What procedures and parameters are established for all aspects of data/information collection necessary to support archiving data collected as part of the evaluation? *Examples include informed consent, data maintenance, de-identifying data procedures.*	All program data, including consent forms, service data, and surveys, will be maintained securely during the project period. Once the evaluation / research period is over, and data analysis and reporting are complete, MER will de-identify the data and make it available for sharing with other researchers or on federal data sites as appropriate. To facilitate this, MER has developed and will be implementing a data archiving plan for this project (as it does for all such major evaluation efforts). MER's Data Archiving plan/processes includes the steps of: 1. Inventorying and determining which data must be archived, and which data will not be archived or shared

	(such as personally identifiable information of participants).

2. Ensuring that all required data elements are maintained in a way that is compliant with OFA and OPRE guidance and regulations and best practices. This includes the creation of supporting documentation such as a code book which makes the data useable by future researchers.

3. Developing an all-inclusive archive policy for the specific data needs, which ensures that archiving is achievable and manageable, and will include the specifics on the duration of data storage (5 years), benchmarks for archiving data, the variety of media to be used for storage – both Dropbox long term storage and physical storage on secured flash drives, and rules for who will have access to the data, including Dropbox controls and physical storage in a locked and fireproof storage facility.

4. Proactively protecting the data archive's integrity by selecting the Dropbox Enterprise company to warehouse the data and by using the protections described above, including the active scanning of stored data. Dropbox also includes search and discovery functions, automated back-up, and total encryption. |
Describe how the collection methods for all types of proposed data collection will support the archiving and transfer of each type.	These systems support the data archiving function and facilitate the sharing of data, when appropriate, because the data collected will be combined from secure database systems and transformed into a tab delimited data sheet which can then be exported into a variety of formats including Excel.
How will consent form language represent plans to store data for sharing and/or transferring to other researchers?	The MER consent forms and processes clearly explain the use and long-term storage of data and the protection of personal and identifiable information to participants (and the de-dentification of data kept long term).
Describe methods of data storage that will support archiving and/or transferring data.	MER's Data storage plan includes both physical and SAS (software as service) processes for long term storage and transfer. As the data set is being built, Qualtrics and Dropbox will be the primary and redundant systems for data storage. Once the all the data has been

	collected and the dataset has been securely and completely de-identified, all data will be removed from Qualtrics and a long-term storage file with limited access will be maintained on Dropbox. In addition, physical download of the data will be made using multiple formats and copies (flash drive and CD or other media available at the time) and this data will be stored securely in two locked locations – a bank vault and MER headquarters protected in a safe behind multiple locks with limited access.
	When transfer is appropriate, either a secure internet-based file transfer can be used, or a delivery of a physically stored data set can be done.
Explain how data and analyses file construction and documentation will support data archiving and/or transferring.	As part of the data analysis and cleanup process, a code book will be constructed that describes the data fields and how they were collected or constructed. Supporting documentation will be stored with the data sets for easy transfer when appropriate.

3. Dissemination

Briefly describe the planned dissemination efforts associated with the local evaluation, including any dissemination that will occur while the evaluation is ongoing (rather than after the evaluation is completed).

AFC and MER are planning on the publication and presentation of the findings of this research effort, both in a descriptive manner in the project's early years and regarding outcomes and impacts over time. MER will produce at least one manuscript for publication and make two professional presentations on the program and its findings. AFC and MER will use the most appropriate and robust analytical method possible in reporting results, including in the final report for the project to OFA, and to making the de-identified dataset available to the public after the conclusion of this project. In addition, we will make use of the HMRF network and conferences to share information with the field, and we will empower Action for Children to use evaluation findings in communications with other partners, funders, and their community.