Fatherhood
Research
& Practice
Network

www.frpn.org

Achieving High Response Rates and Dealing with Missing Data in Fatherhood Evaluations

June 13, 2017

Who is FRPN?

Jay Fagan, Ph.D.
Temple University
FRPN Co-Director

Rebecca Kaufman, MSW
FRPN Senior Research Coordinator

Jessica Pearson, Ph.D.
Center for Policy Research
FRPN Co-Director

Nancy Thoennes, Ph.D.
Center for Policy Research

Overview of FRPN

- Five-year, $4.8 million cooperative agreement to Temple University/CPR

- Funding by U.S. DHHS, ACF, Office of Planning, Research and Evaluation, October 2013-September 2018

- Targets fatherhood researchers & programs serving low-income fathers (OFA and non-OFA grantees, state fatherhood commissions, CBOs, programs funded by TANF, child welfare & child support)

Fatherhood Research And Practice Network

#1 Promote Rigorous Evaluation

#2 Build Evaluation Capacity

#3 Disseminate Information

Contact Us

- Jay Fagan, PhD, FRPN Co-Director
  - jfagan@temple.edu
  - (215) 204-1288

- Jessica Pearson, PhD, FRPN Co-Director
  - jspearson@centerforpolicyresearch.org
  - (303) 837-1555

- Rebecca Kaufman, MSW, FRPN Coordinator
  - rebecca.kaufman@temple.edu
  - (215) 204-5706

Why a Webinar on Response Rates and Missing Data?

- Fatherhood program evaluations often suffer from low response rates.

- Mathematica researchers achieved a response rate of 70% in the P.A.C.T. evaluation.

- To improve response rates, programs should collect certain types of contact information & consents, and maintain regular contact with program participants.

- Good response rates depend on the strength of the field effort including levels and methods of outreach.

- The utility of various statistical techniques to handle missing data depends on a lot of things: the amount of missing data, its randomness, overall sample size, and the nature of the measures.

Today's Speakers

Cleo Jacobs Johnson, PhD is a Senior Survey Researcher at Mathematica Policy Research.

Shawn E. Marsh, Associate Director of Survey Research, Mathematica Policy Research

Quinn Moore is a senior researcher at Mathematica Policy Research.

Participant Retention in Hard-to-Reach Populations:

Strategies for Outreach

Fatherhood Research & Practice Network Webinar

June 13, 2017

Shawn Marsh • Cleo Jacobs Johnson

Overview

- **Current climate for contacting participants**

- **Three areas of focus to reach and retain participants**

- **A practical example from a recent project**
  - **Parents and Children Together (PACT) evaluation of Responsible Fatherhood (RF) and Healthy Marriage (HM) programs**
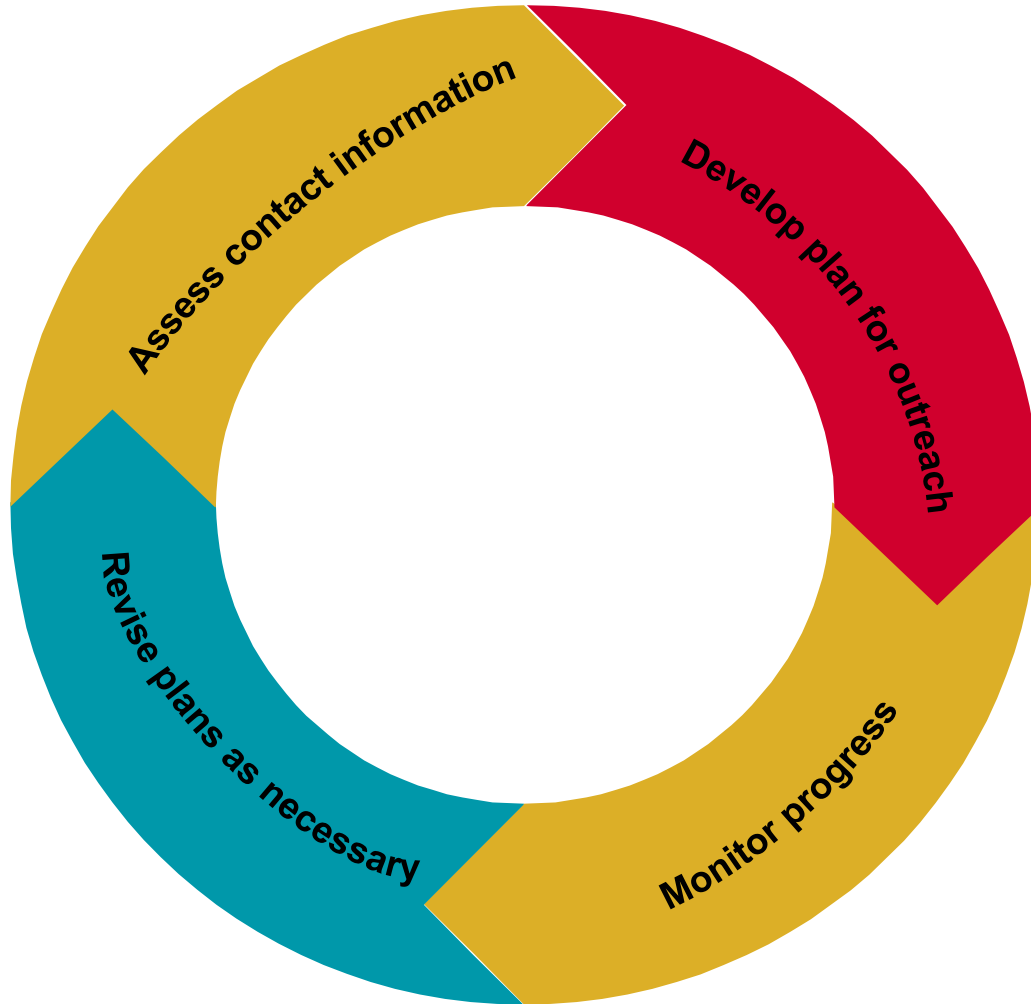

PACT
☑ HM: 88%
☑ RF: 72%
Thanks for your hard work!

MATHEMATICA
Policy Research

Current climate for contacting participants

- **Proliferation of cell phones**

- **Research fatigue**

- **Privacy concerns**

- **Continually changing life circumstances**

- **Reluctance to engage with perceived authorities**

MATHEMATICA
Policy Research

Overcoming the Challenges

MATHEMATICA
Policy Research

Key steps for outreach

A circular cycle diagram showing four steps for outreach:
- Develop plan for outreach
- Monitor progress
- Revise plans as necessary
- Assess contact information

MATHEMATICA
Policy Research

PACT
Parents and Children Together

Assessing Contact Information

- **Assess contact information for sample members:**
  - What types of information do you have?
  - How many different contacts do you have?
  - How complete is the information for each contact?

- **Train staff on the importance of getting thorough contact information**
  - Staff members can be reluctant to ask about information they view as sensitive (e.g., SSN, DOB)
  - All of it is PII, so it's important to stress the importance of handling it with care, but not be scared collecting it
  - It's useful to explain how you're going to protect their information

Planning for Outreach

- **Maintaining contact requires well-thought out planning**
  - Determine how you plan to use contact information
  - Determine the frequency of contact and when contact will occur
  - Decide who you are contacting; is it the sample member and/or their relatives and friends and what are you telling them
  - Plan to include special populations (e.g., military, prison) that may require extra effort, but are not prohibitive

- **Re-evaluate your plan periodically**
  - Review the results of your contacts and update your plan
  - Make sure you aren't overdoing the contacts

MATHEMATICA
Policy Research

Monitoring Progress

- **Identify a leader for locating and retention**
  - Develop scripts and guides
  - Ensure that participants are called back as requested

- **Create tracking system**
  - Make sure that new contact information is updated
  - Use new contact information (e.g., addresses, phone numbers, email addresses)

- **Follow up on leads in a timely manner**

MATHEMATICA
Policy Research

Description of the PACT Study

- **Funding provided by the Administration for Children and Families (ACF)**

- **Study objectives:**

  1. To understand how PACT programs were designed and implemented

  2. To learn how participants view and carry out their roles as parents, providers, and partners

  3. To evaluate whether selected RF and HM programs improve outcomes for enrolled fathers and couples

MATHEMATICA
Policy Research

PACT 12-Month Follow-up Challenges

- **Following up with a hard-to-reach population is difficult**

- **Evaluation sites recruited from homeless shelters and halfway houses**

- **Participants experienced significant life changes between baseline and follow-up, including incarceration and homelessness**

Assessing Contact Information

- **On PACT,** *we* **collected the contact information**
  - **We knew what we had, which was an advantage, but there were still holes**
  - **We had a number of cases without a valid telephone number**
  - **We analyzed contact information and figured out that a lot of people were in homeless shelters**

MATHEMATICA
Policy Research

We asked for detailed contact information

Participants

- SSN

- Physical address

- Home telephone

- Cell phone

- Features of cell phone plan

- Permission to text

- Email address (primary and secondary)

- Social media account names (Facebook, Twitter, MySpace, others)

3 family members or friends

- Physical address

- Relationship to participant

- Home telephone

- Cell phone

- Work phone

- Email address (primary and secondary)

- Social media account names (Facebook, Twitter, MySpace, others)

MATHEMATICA
Policy Research

PACT
Parents and Children Together

Homeless shelter address provided by multiple fathers

1621 First Street			Saint Louis	MO	63102-____
1621 N 1 st street			Saint Louis	MO	63102-____
1621 N 1st St			Saint Louis	MO	63102-____
1621 N 1st St			Saint Louis	MO	63102-____
1621 N 1st St			Saint Louis	MO	63102-____
1621 N 1st St			Saint Louis	MO	63103-____
1621 N 1st St			Saint Louis	MO	63102-____
1621 N 1st St			Saint Louis	MO	63102-____
1621 N 1st St			Saint Louis	MO	63102-____
1621 N 1st Street			Saint Louis	MO	63102-____
1621 N 1st Street			Saint Louis	MO	63102-____
1621 N 1st Street			Saint Louis	MO	63102-____
1621 N 1st Street			Saint Louis	MO	63102-____
1621 N 1st Street			Saint Louis	MO	63102-____
1621 N 1st Street			Saint Louis	MO	63103-____
1621 N 1st Street			Saint Louis	MO	63102-____
1621 N First St			Saint Louis	MO	63102-____
1621 N. 1st St			Saint Louis	MO	63120-____
1621 N. 1st St			Saint Louis	MO	63103-____
1621 N. First St.			Saint Louis	MO	63113-____
1621 N. First St.			Saint Louis	MO	63102-____

Developing a Data Collection Plan

- **We had a plan for outreach that included:**
    - **Notification strategy**
    - **Database searches**
    - **Plans for transitioning cases from one mode to another**

MATHEMATICA
Policy Research

Multi-step notification strategy

Release Month	TLO Batch	Release Date and 1st Mailing Batch	2nd Mailing Batch	3rd Mailing Batch	4th Mailing Batch	Locating and Refusal Letters (1st Thursday)	Locating and Refusal Letters (2nd Thursday)	Locating and Refusal Letters (3rd Thursday)	Locating and Refusal Letters (4th Thursday)	Locating and Refusal Letters (5th Thursday *if applicable*)
Release	Run On	Mail-Out Date	Send-Out Date	Send-Out Date	Send-Out Date	Mail-Out Date	Mail-Out Date	Mail-Out Date	Mail-Out Date	Mail-Out Date
January 2015	Dec-22-2014	Dec-29-2014	Jan-05-2015	Jan-12-2015	Jan-19-2015					
February 2015	Jan-19-2015	Jan-26-2015	Feb-02-2015	Feb-09-2015	Feb-16-2015					
March 2015	Feb-16-2015	Feb-23-2015	Mar-02-2015	Mar-09-2015	Mar-16-2015	Mar-05-2015	Mar-12-2015	Mar-19-2015	Mar-26-2015	
April 2015	Mar-16-2015	Mar-23-2015	Mar-30-2015	Apr-6-2015	Apr-13-2015	Apr-02-2015	Apr-09-2015	Apr-16-2015	Apr-23-2015	Apr-30-2015
May 2015	Apr-20-2015	Apr-27-2015	May-4-2015	May-11-2015	May-18-2015	May-07-2015	May-14-2015	May-21-2015	May-28-2015	
June 2015	May-18-2015	May-25-2015	Jun-1-2015	Jun-8-2015	Jun-15-2015	Jun-04-2015	Jun-11-2015	Jun-18-2015	Jun-25-2015	
July 2015	Jun-22-2015	Jun-29-2015	Jul-6-2015	Jul-13-2015	Jul-20-2015	Jul-02-2015	Jul-09-2015	Jul-16-2015	Jul-23-2015	Jul-30-2015
August 2015	Jul-20-2015	Jul-27-2015	Aug-3-2015	Aug-10-2015	Aug-17-2015	Aug-6-2015	Aug-13-2015	Aug-20-2015	Aug-27-2015	
September 2015	Aug-17-2015	Aug-24-2015	Aug-31-2015	Sep-7-2015	Sep-14-2015	Sep-03-2015	Sep-10-2015	Sep-17-2015	Sep-24-2015	
October 2015	Sep-21-2015	Sep-28-2015	Oct-5-2015	Oct-12-2015	Oct-19-2015	Oct-01-2015	Oct-08-2015	Oct-15-2015	Oct-22-2015	Oct-29-2015
November 2015	Oct-19-2015	Oct-26-2015	Nov-2-2015	Nov-9-2015	Nov-16-2015	Nov-5-2015	Nov-12-2015	Nov-19-2015	Nov-26-2015	
December 2015	Nov-16-2015	Nov-23-2015	Nov-30-2015	Dec-7-2015	Dec-14-2015	Dec-03-2015	Dec-10-2015	Dec-17-2015	Dec-24-2015	Dec-31-2015

Monitoring Progress

- **We had a task leader focused specifically on locating**

  – **We had an internal tracking system**

  – **We had standardized reports to manage the process**

MATHEMATICA
Policy Research

Sample monitoring report

Category and Current Status (Based On LOCATING STATUS)		07/26/15		Cumulative			Calls
		N Cases	% Column	N Cases	% Column	% Category	Mean
LOCATED	1860 Review by supervisor needed for field	6	15.79	66	1.95	3.05	7.0
	1870 Review completed, ready for field	0	0.00	26	0.77	1.20	0.0
	1873 Ready for field- evasive	0	0.00	1	0.03	0.05	1.0
	1874 Ready for field- refusal	0	0.00	6	0.18	0.28	0.2
	1880 Sent to field	0	0.00	40	1.18	1.85	0.1
	1890 Phone Located - return to CATI	1	2.63	1151	33.98	53.24	3.9
	1891 Address Located	0	0.00	871	25.72	40.29	1.2
	1899 New telephone number, not yet attempted	0	0.00	1	0.03	0.05	0.0
	SUBTOTAL	7	18.42	2162	63.83	100.00	2.8
TO LOCATING	1510 Not-in service / Not a working number	2	5.26	2	0.06	0.42	4.0
	1530 Wrong number / No such person	5	13.16	13	0.38	2.75	9.1
	1538 Threshold reached, send to locating	0	0.00	7	0.21	1.48	19.9
	1540 Moved out of area	0	0.00	1	0.03	0.21	0.0
	1541 Incarcerated Respondent, send to locating	0	0.00	8	0.24	1.69	1.6
	1542 Pre-Locating	0	0.00	16	0.47	3.38	0.1
	1566 Return to Locating from field	0	0.00	165	4.87	34.88	0.2
	1581 New phone found in field	1	2.63	1	0.03	0.21	0.0
	2590 Final unlocatable by phone center	0	0.00	161	4.75	34.04	0.0
	2591 Final unlocatable by field staff	1	2.63	99	2.92	20.93	0.1
	SUBTOTAL	9	23.68	473	13.97	100.00	0.7
ACTIVE IN LOCATING	1903 Active locating - call night	0	0.00	24	0.71	3.24	7.6
	1904 Active locating - call weekday	9	23.68	27	0.80	3.64	6.2
	1905 Active locating - call weekend	0	0.00	9	0.27	1.21	9.2
	1909 Needs final locating status	1	2.63	42	1.24	5.67	0.2
	1915 Internet searching, businesses / professionals	0	0.00	3	0.09	0.40	0.0

MATHEMATICA
Policy Research

Conclusions

- **Locating and contacting participants is more complicated than people think, but it can be scaled to the resources you have available**

- **The most important things are to assess, plan and monitor**

- **If you have the resources, hiring experts can be worth the money**

MATHEMATICA
Policy Research

For More Information

- **Shawn Marsh**
  - **smarsh@mathematica-mpr.com**

- **Cleo Jacobs Johnson**
  - **cjacobs@mathematica-mpr.com**

MATHEMATICA
Policy Research

Approaches to dealing with nonresponse

June 13, 2017

Quinn Moore

Sample attrition can bias impact estimates

- **Well-designed study has research groups that are similar other than whether offered intervention**

- **With sample attrition, composition of research groups changes**

  – **Nonrespondents are not randomly selected**

  – **Response rates might be lower for important demographic groups**

  – **Response rates might be lower for control group**

- **Biased impacts = wrong answers!**

MATHEMATICA
Policy Research

When to worry about attrition bias?

- **Attrition bias risk increases based on combination of overall attrition and differential attrition**

  – **Overall attrition: attrition for the entire sample**

  – **Differential attrition: difference in attrition between research groups**

- **Researchers and federal systematic reviews have developed standards for risk of attrition bias**

  – **ACF: SAFER, ESER**

  – **DoL: CLEAR**

  – **DoE: What Works Clearinghouse**

MATHEMATICA
Policy Research

Examples of attrition bias risk assessment

- **Examples of combinations of overall and differential attrition considered low risk of bias**

Overall attrition	Differential attrition
10	6.3
20	5.4
30	4.1
40	2.6
50	1.0

- **If fail attrition test, need to show equivalence on baseline characteristics selected ahead of time**

MATHEMATICA
Policy Research

Conducting nonresponse analysis

- **What is the overall response rate? What is the difference in the response rates for the research groups?**

  – **Assess attrition bias risk**

- **How do the baseline characteristics of respondents and nonrespondents compare, overall and by research group?**

  – **Look for selection into survey response and whether selection varies for research groups**

- **How do the baseline characteristics of respondents compare by research group?**

  – **Speaks to validity of impact estimates for respondents (the analysis sample)**

MATHEMATICA
Policy Research

Strategies for Dealing with Nonresponse

MATHEMATICA
Policy Research

Strategies for mitigating attrition bias: Prevention

- **Prevention is by far the best solution**

- **If attrition is high, researchers might be skeptical of findings regardless of statistical treatment**
  - Systematic reviews assign high risk based on attrition only

- **Maintain high survey response**
  - Monitor overall and differential attrition during data collection
  - Strategies discussed in earlier talk

- **Use all available sample in analysis**

MATHEMATICA
Policy Research

Reduce attrition by avoiding truncated outcomes

- **A different type of attrition: When outcomes are not defined for the full sample**

  – **These are called truncated outcomes**

- **Particularly common in fatherhood and family research**

- **Examples:**

  – **Relationship quality is defined only if in contact with child/CP**

  – **Hourly wage is defined only if have job**

  – **Employment benefit defined only if have job**

MATHEMATICA
Policy Research

Truncation is analogous to nonresponse: Missing data leads to bias

	Impact on Contact	Impact on Relationship Quality
True Impact	+	○
Observed Impact in Truncated Sample		

MATHEMATICA
Policy Research

Truncation is analogous to nonresponse: Missing data leads to bias

	Impact on Contact	Impact on Relationship Quality
True Impact	+	○
Observed Impact in Truncated Sample	+	

MATHEMATICA
Policy Research

Truncation is analogous to nonresponse: Missing data leads to bias

	Impact on Contact	Impact on Relationship Quality
True Impact	+	○
Observed Impact in Truncated Sample	+	–

MATHEMATICA
Policy Research

Strategies for dealing with truncation

- **Define outcomes for all sample members when possible**
  - Total earnings rather than hourly wage

- **Use binary outcomes where appropriate**
  - In a high quality relationship rather than relationship quality
  - Employed in a job offering benefits

- **If necessary to use truncated outcome**
  - Assess risk of attrition bias for truncated sample

MATHEMATICA
Policy Research

Strategies for mitigating attrition bias: Statistical approaches

- **Case deletion**

- **Regression adjustment**

- **Single imputation**

- **Multiple imputation**

- **Nonresponse weights**

MATHEMATICA
Policy Research

Strategies for mitigating attrition bias: Case deletion

- **Delete those who have missing data—essentially do nothing**

- **Assumes that nonrespondents are the same as respondents, on average**

- **Easy to implement and transparent**

- **If risk of attrition bias is low, this is defensible**

- **Some simulation evidence that this approach is appropriate for evaluations under many circumstances (Puma et al. 2009)**

  – **Improved by adding regression adjustment**

MATHEMATICA
Policy Research

All statistical approaches other than case deletion require baseline data

- **Need to know who the nonrespondents were and how they differed from respondents**

- **Looking for factors that are associated with survey response**

  – **Demographic characteristics**

  – **Data related to data collection: Have cell phone, number of contacts, housing stability, etc.**

MATHEMATICA
Policy Research

Statistical approaches to attrition: Regression adjustment

- **Estimate impacts while controlling for factors associated with nonresponse in a regression**

- **Accounts for differences between treatment and control respondents**

  – **Leads to unbiased impacts for respondents**

- **Does not help if impacts are different for respondents and nonrespondents**

MATHEMATICA
Policy Research

Regression adjustment can be combined with other approaches

- Improves precision of impact estimates

- Education simulations suggest that controlling for the baseline version of the outcome reduces bias by 50% in samples with high attrition

- Recommended for studies with high attrition

MATHEMATICA
Policy Research

Statistical approaches to attrition: Single imputation

- **Fill in missing values for nonrespondents**

- **Approaches include**
  - **Mean fill**
  - **Hot deck**
  - **Regression fill**

- **Overstates precision of estimates because doesn't take into account uncertainty associated with imputation**

- **Can lead to biased estimates**

- **Not recommended**

MATHEMATICA
Policy Research

Statistical approaches to attrition: Multiple imputation

- **Increasingly popular as computing power and software availability have advanced**
  - Stata and SAS have modules to deal with multiply imputed data

- **Fill in values that are missing due to nonresponse and account for uncertainty associated with imputation**

MATHEMATICA
Policy Research

Multiple imputation steps

- **Generate data set with no missing values using available information**

  - **Using baseline data to make an educated guess at missing value**
  - **Predict values based on regression and add random term**
  - **Rerun regressions and get new predicted values**
  - **Repeat until stable**

- **Do this several times**

  - **5 to 30 complete data sets**

- **Calculate impacts**

  - **Impact estimate is the average of the impact across the imputed data sets**
  - **Standard error of estimate accounts for uncertainty of imputation**

MATHEMATICA
Policy Research

Multiple imputation implementation

- **Imputation must be conducted separately by research group**

- **Imputation must include covariates used in impact estimation**

- **Imputation should include factors related to response status**

MATHEMATICA
Policy Research

When is multiple imputation most useful?

- **Severity of attrition problem**
  - **If very severe, no approach will be convincing**

- **Quality of baseline data**

- **Strength of relationship between measured baseline factors and outcomes**

MATHEMATICA
Policy Research

Challenges with multiple imputation

- **Implementation is not always as straightforward as advertised**

- **More challenging with more complicated data**
  - **Truncated variables**
  - **Clustered and other complex designs**

- **Some researchers object to imputing outcomes**

MATHEMATICA
Policy Research

Statistical approaches to attrition: Nonresponse weights

- **Find respondents that look the most like nonrespondents and give them bigger weights**

- **Weights are inversely proportional to the likelihood of survey response**

MATHEMATICA
Policy Research

Calculating nonresponse weights

- **Software does not "automatically" generate weights**

- **Propensity score approach**
  - **Logistic regression predicting probability of response**
  - **Weight is inverse of predicted value**
  - **Easy to calculate, but can have large values**

- **Weight class approach**
  - **Divide sample into groups**
  - **Calculate response rate for each group**
  - **Calculate weight based on probability of response**

- **Distributions of weights need to be examined carefully**
  - **Extreme weight values often trimmed**
  - **Caution is warranted**

MATHEMATICA
Policy Research

When are nonresponse weights most useful?

- **Severity of attrition problem**

- **Quality of baseline data**

- **Strength of relationship between measured baseline factors and survey response**

MATHEMATICA
Policy Research

Approach for PACT and BSF

- **Survey nonresponse: weights**
  - Separate for mothers, fathers, couples

- **Item nonresponse: multiple imputation**
  - Allows taking advantage of responses to other related survey items

- **Allows for robustness checks**
  - Plain case deletion, case deletion with regression adjustment, multiple imputation with and without weights

- **Similar results across methods**
  - Consistent with low attrition bias risk

How to plan for nonresponse at different project stages?

- **Data collection**
  - Strategies for achieving high response
  - Monitor overall and differential attrition

- **Analysis planning**
  - Select baseline variables to include in nonresponse/equivalence analysis
  - Define outcomes to maximize inclusiveness of sample; avoid truncated outcomes when possible
  - Determine strategy for dealing with nonresponse (case deletion with regression adjustment, MI, weighting)

- **Analysis and reporting**
  - Conduct nonresponse analysis
  - Consider implications of nonresponse in interpreting findings
  - Control for baseline version of outcomes

Questions

- **Is there a rule of thumb in using MI vs simpler less sophisticated imputation techniques depending on how much data is missing on a given variable?**

- **Are there limits to the types of models that can be used when producing estimates using MI for missing data?**

- **Is it ever OK to keep a case in a study when pre-test data are missing?**

- **What type of survey items are appropriate to impute? Does this vary with item nonresponse rates?**

MATHEMATICA
Policy Research