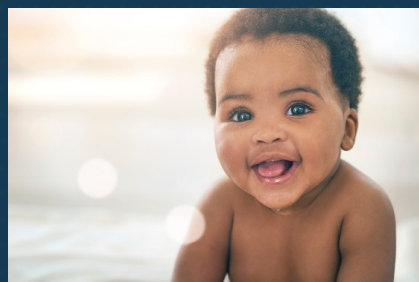# MEASURING NONRESPONSE BIAS IN THE PREGNANCY RISK ASSESSMENT MONITORING SYSTEM (PRAMS), 2019

HOLLY SHULMAN, MA; PHIL HASTINGS, PHD*; RUBEN SMITH, PHD; LEE WARNER, PHD

THE FINDINGS AND CONCLUSIONS IN THIS REPORT ARE THOSE OF THE AUTHORS AND DO NOT NECESSARILY REPRESENT THE OFFICIAL POSITION OF THE CENTERS FOR DISEASE CONTROL AND PREVENTION.

Centers for Disease Control and Prevention

National Center for Chronic Disease Prevention and Health Promotion

Division of Reproductive Health/ *Far Harbor LLC

APPROACH

- Jurisdiction vital records birth file serves as sampling frame and source of population information

- Linked birth file with PRAMS sample for 47 PRAMS sites (3 sites did not grant permission for access to their full birth file) to compute:

  - Population values

  - Estimates of the population values from the full PRAMS sample (using sampling weights only)

  - Estimates of the population values from PRAMS respondents (using analysis weights)

BIRTH CERTIFICATE VARIABLES EXAMINED

Health Behaviors	Medical	Demographic/SES
Adequate prenatal care (PNC)	Gestational diabetes	Medicaid as payment source for delivery
Smoking before pregnancy	Gestational high blood pressure (hypertension)	WIC participant
Smoking during pregnancy	C-Section delivery	Previous live birth
Breastfed in hospital	Infertility treatment	
	Pre-pregnancy Body Mass Index (BMI) - Normal	

Note: WIC is the Special Supplemental Nutrition Program for Women, Infants, and Children

# COMPARISON OF RESPONDENT ESTIMATES TO POPULATION VALUES

- Produce 95% confidence interval (CI) for estimates of each indicator from PRAMS respondents in each jurisdiction

- Identify instances where 95% CI excludes true population value

- Compute bias and absolute relative bias

$$Bias_{NR}(\overline{Y}_r) = \overline{Y}_r - \overline{Y}_f$$

$$Absolute\ Relative\ Bias_{NR}(\overline{Y}_r) = \left| \frac{100\% * (\overline{Y}_r - \overline{Y}_f)}{\overline{Y}_f} \right|$$

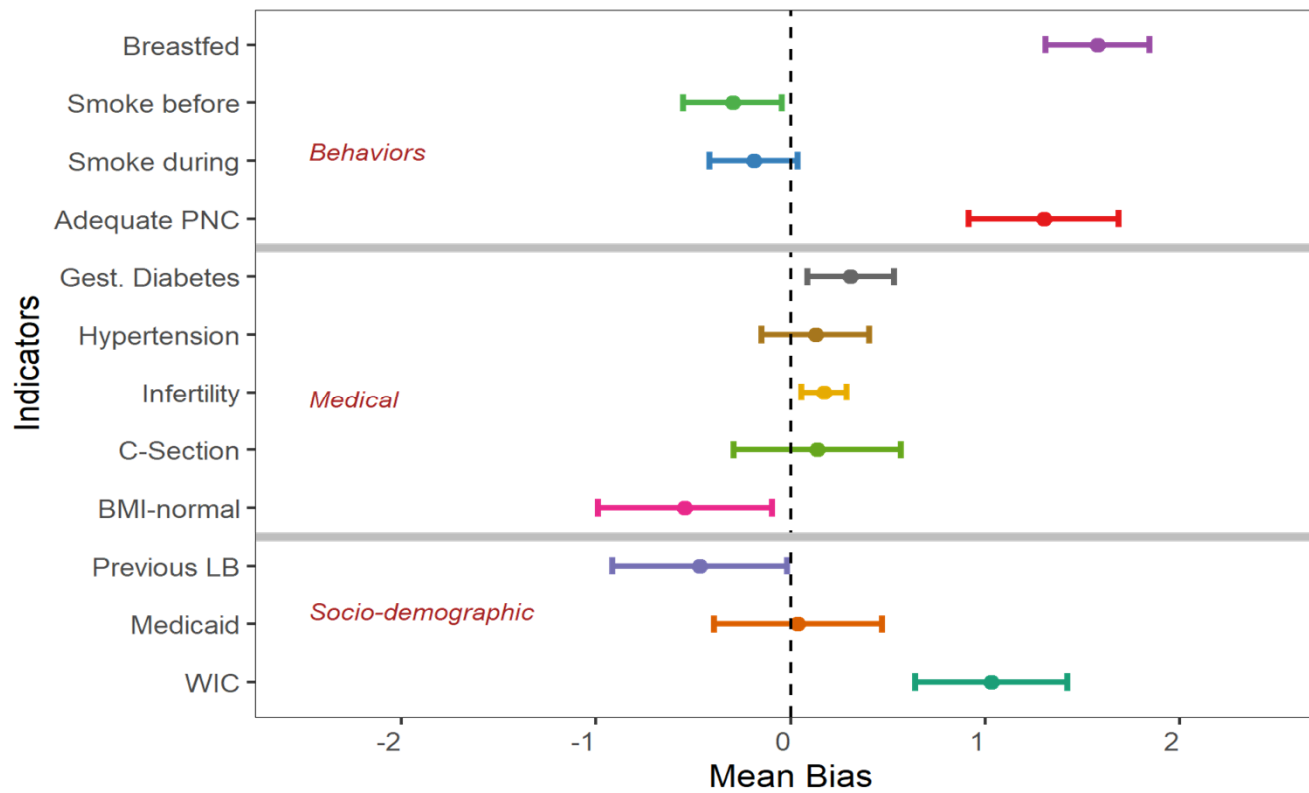where $\overline{Y}_f$ is the frame value and $\overline{Y}_r$ is the respondent weighted estimate

COMPARISON OF RESPONDENT ESTIMATES TO POPULATION VALUES - NUMBER OF SIGNIFICANT DIFFERENCES (47 SITES)
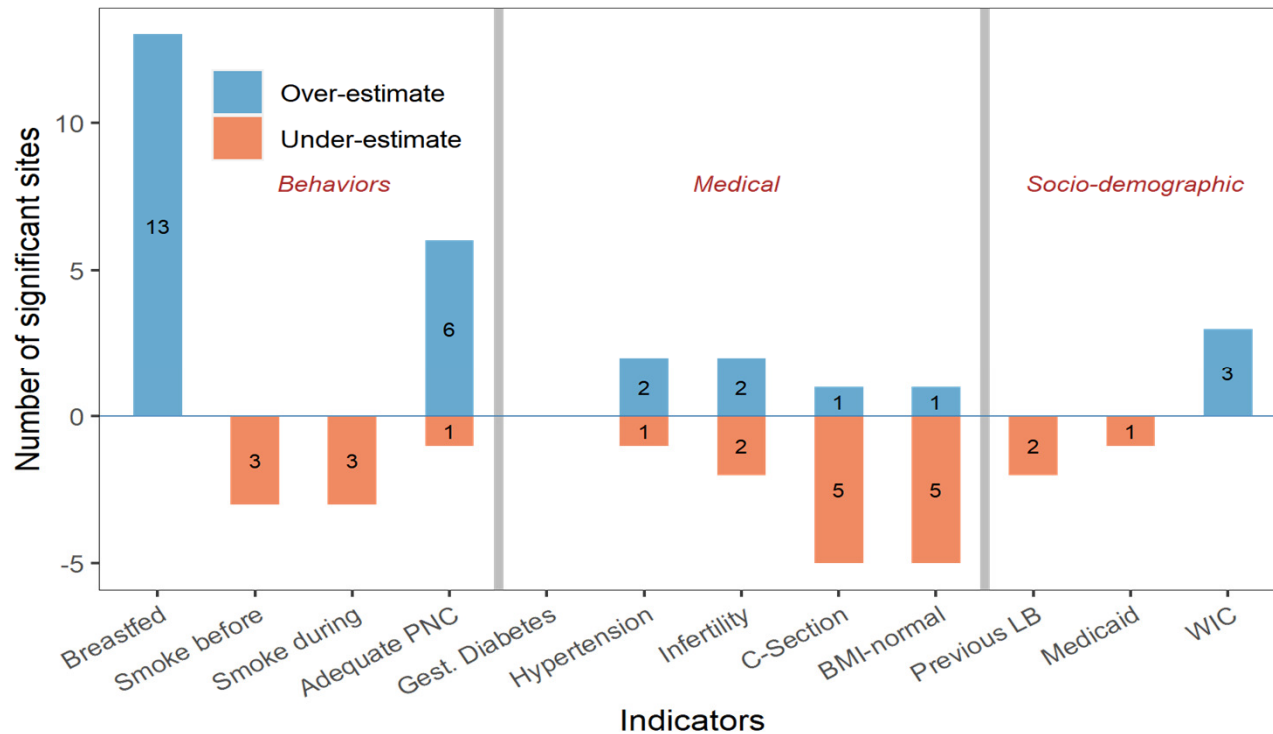
MEAN BIAS BY INDICATOR

- Averaged bias over all 47 sites to produce the mean bias for each indicator

- For each indicator, conducted a fixed-effects meta-analysis with each site treated as a study to account for different sample sizes across sites

  - Effect size measure for the meta-analysis was the (logged) risk ratio, comparing "risk" (prevalence) of reporting in the weighted PRAMS sample to prevalence in the population at each site

  - Allows us to estimate a confidence interval around the mean bias

MEAN BIAS ACROSS SITES BY INDICATOR

IMPACT OF RESPONSE RATES ON BIAS

- Calculate mean absolute bias (MAB) across all 12 indicators for each site

$$Mean\ Absolute\ Bias_{NR}(\overline{Y}_r) = \sum_{i=1}^{12} |\overline{Y}_{ri} - \overline{Y}_{fi}| \ / \ 12$$

where $\overline{Y}_{fi}$ is the frame value for indicator i and $\overline{Y}_{ri}$ is the respondent weighted estimate for indicator i

- Compute least squares regression line for MAB as predicted by response rate

RELATIONSHIP BETWEEN MEAN ABSOLUTE BIAS AND RESPONSE RATES
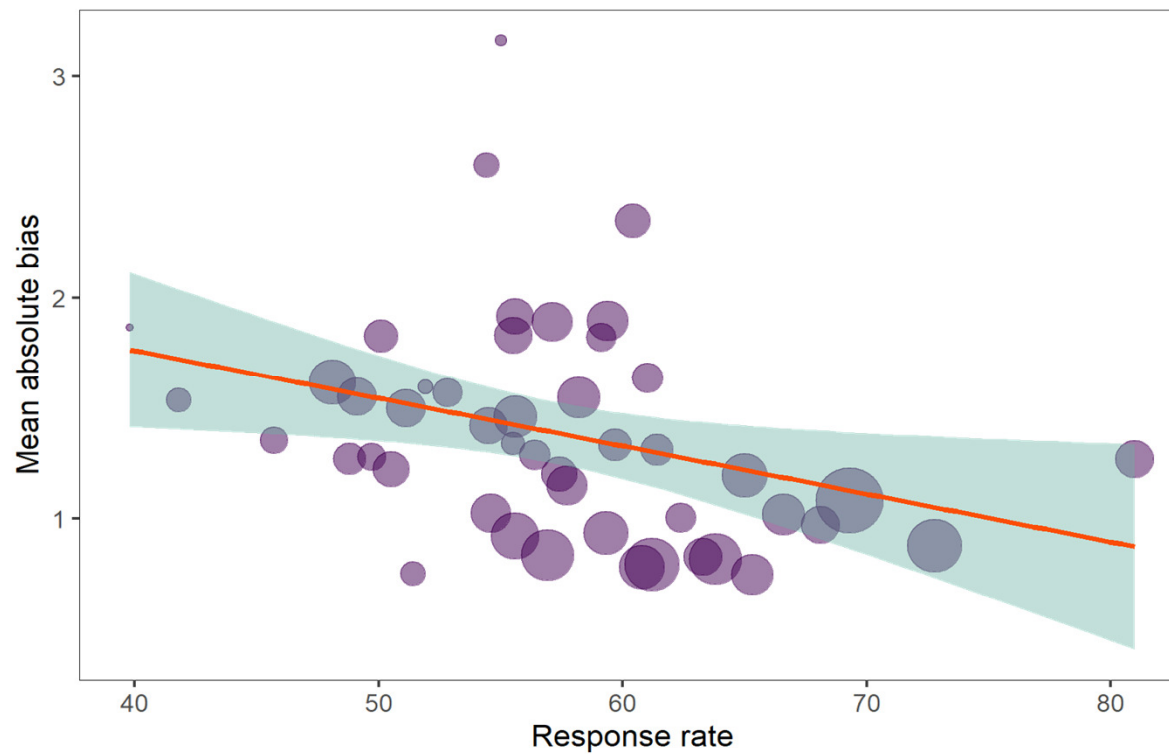
Least Squares Regression Line (b = -0.021, SE = 0.009, t = =2.35, p = .023)

Correlation between response rate and bias = -0.33

COMPARISON OF RESPONDENT ESTIMATES, FULL SAMPLE ESTIMATES, AND POPULATION VALUES

- Calculate bias and absolute bias for both the full sample estimates (i.e. 100% response) and respondent estimates for three indicators: breastfed in hospital, gestational diabetes, and WIC participant

- Compute mean bias and mean absolute bias over all 47 jurisdictions

$$Mean\ Bias_{NR}(\overline{Y}_{fs}) = \sum_{i=1}^{47}(\overline{Y}_{fsi} - \overline{Y}_{fi}) / 47$$

$$Mean\ Absolute\ Bias_{NR}(\overline{Y}_{fs}) = \sum_{i=1}^{47}|\overline{Y}_{fsi} - \overline{Y}_{fi}| / 47$$

where $\overline{Y}_{fi}$ is the frame value for indicator i and $\overline{Y}_{fsi}$ is the full sample weighted estimate for indicator i

- Similar computation for respondent sample weighted estimate

MEAN BIAS AND MEAN ABSOLUTE BIAS FOR FULL SAMPLE AND RESPONDENT ESTIMATES (47 SITES)
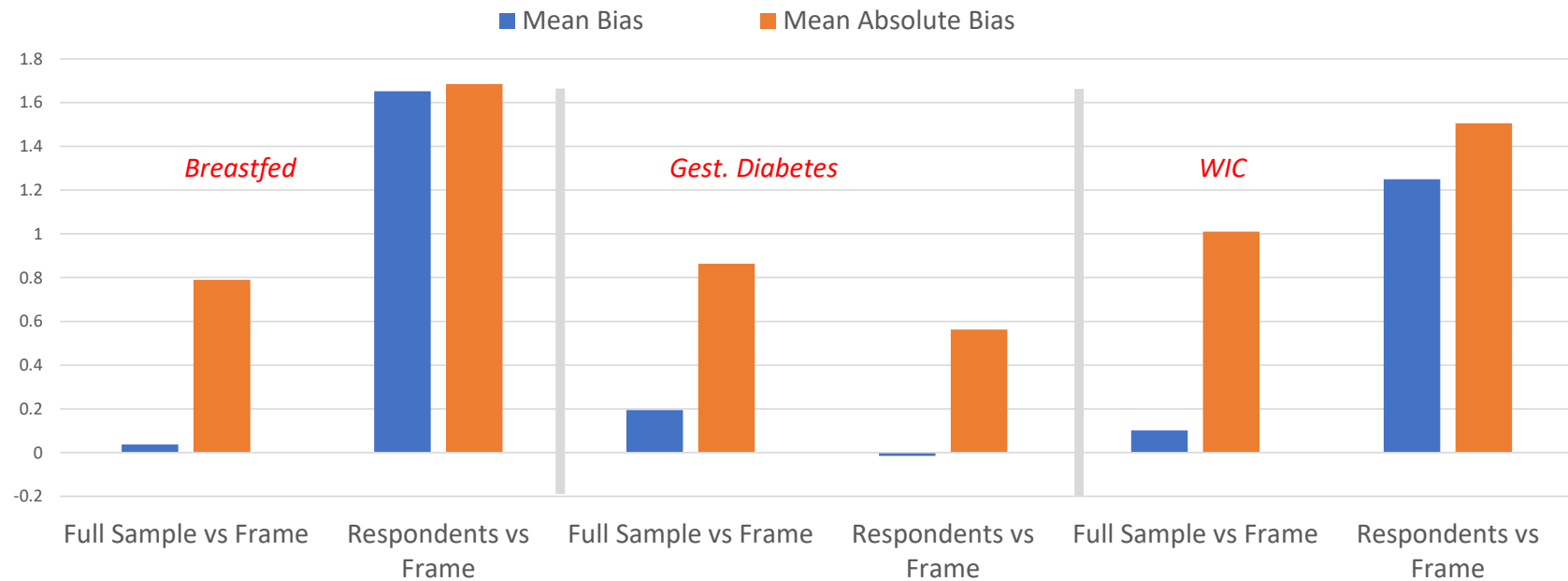
RESULTS

- Actual bias observed in PRAMS was relatively small

  - Highest mean absolute bias was 1.68 percentage points for breastfeeding

  - 9 of 12 indicators had mean absolute bias <1 percentage point

- Behavioral indicators had highest levels of bias; demographic/socio-economic indicators had the lowest

- Positive behaviors were over-estimated; unhealthy behaviors were under-estimated

- Weak correlation (-0.33) between response rate and bias

- Mean bias of full sample estimates was very small, as would be expected from a series of independent random samples

CONCLUSIONS

- Observed levels of bias are acceptable for most uses of PRAMS data

- Behavioral indicators might be more susceptible to bias – social desirability or other individual- or group-level influences may impact reporting of behavioral indicators which are not amenable to weighting adjustments

- Bias varies across indicators (even within categories); must be examined at the indicator level

STRENGTHS AND LIMITATIONS

- Strengths

  - Access to complete, population-level data for a broad range of indicators

  - Standard methodology use by PRAMS sites allows for analysis of impact of response rate on bias while controlling for other factors associated with bias

  - Overlap of auxiliary variables on frame (medical and behavioral) with survey topics

- Limitations

  - Limited set of indicators collected on birth certificates

- Caveats

  - Results cannot be generalized beyond study population (people who recently delivered a live-born infant)

CONTACT INFORMATION

Holly Shulman, MA
Division of Reproductive Health, CDC
hbs1@cdc.gov

Lee Warner, PhD
Division of Reproductive Health, CDC
dlw7@cdc.gov

Phil Hastings, PhD
Principal, Far Harbor LLC
phil@farharbor.com

Ruben Smith, PhD
Division of Reproductive Health, CDC
eyb4@cdc.gov

PRAMS webpage:
https://www.cdc.gov/prams/