Modeling Total Survey Error in the 2011 National Immunization Survey (NIS): Children and Teens
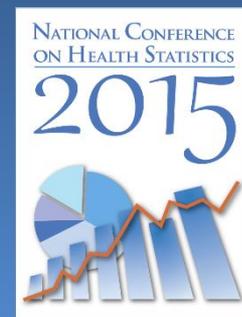
Meena Khare[2], Vicki Pineau[1], Kirk Wolter[1], Ben Skalland[1], and Wei Zeng[1], Carla Black[2], Christina Dorell[2] , and David Yankey[2]

[1]NORC at the University of Chicago
[2]Centers for Disease Control and Prevention

**National Conference on Health Statistics
August 25, 2015**

NIS: An Overview

- Household CATI survey followed by a mailed provider record check study(PRC)
  - Landline RDD sampling frame before 2011 for official estimates
  - Since 2011 Cell RDD samples have been added
  - Household interview screens
  - Children: 19-35 months of age
  - Teens: 13-17 years of age
- Assess and monitor vaccination coverage rates of children and teens in the United States.
- Publish official vaccination estimates at the national, state, and selected local area levels
- Sponsored by the Centers for Disease Control and Prevention

Total Survey Error (TSE)

TSE is the sum of errors from each survey stage and includes:

- Sampling Error
- Non Sampling Error
  - Noncoverage
  - Nonresponse

TSE Analysis can help answer questions:

- What is the size of bias due to noncoverage and nonresponse and what is its impact on estimated vaccination coverage rates?
- What is the impact of different weighting methodologies on the total survey errors?

Comparison of TSE Analysis in the NIS (2010 vs. 2011)

2010

- Assessed bias for multiple NIS children and Teen vaccination coverage estimates
- Considered noncoverage bias potentially introduced by missing the cell-phone-only and zero-bank population along with no phone population

2011

- Repeated TSE to include improvement in population coverage by including a RDD cell phone frame
- Assessed bias under two different dual frame weighting approaches (shrinkage weighting vs. unbiased weighting)
- Assessed nonresponse bias from decreased response rates

4

Universe of Eligible Children

Stage 1: Coverage

Telephone Number in RDD
Dual Sampling Frame
(p_1, μ_1)

No Telephone (p_{1A}, μ_{1A})

Zero-Bank (Landline Only)
(p_{1B}, μ_{1B})

Stage 2: Response

Adequate Provider Data --
Respondents
(p_2, μ_2)

No Adequate Provider Data --
Nonrespondents
$(1-p_2, \mu_{2A})$

Methodology for TSE Analysis (1)

Total Survey Error Analysis Steps:

A. Develop a model describing the survey stages at which component error may enter:

1. *Coverage*
2. *Response (Both the Household interview and the mailed PRC)*

- **Bias due to Nonresponse**

$$q_1 = \hat{\mu}_1 - \mu_1 = \hat{p}_2 q_2 + (1 - \hat{p}_2)q_{2A} - v_2(q_2 - q_{2A})$$

- **Total Bias**

$$q_0 = \hat{\mu}_0 - \mu_0$$
$$= (1 - \hat{p}_{1A} - \hat{p}_{1B})q_1 + \hat{p}_{1A}q_{1A} + \hat{p}_{1B}q_{1A} +$$
$$v_{1A}(q_1 - q_{1A}) + v_{1B}(q_1 - q_{1B})$$

μ denotes the conditional mean of vaccination coverage rates among
children/teens living in a certain household;
p denotes the corresponding probability of living in such households.
v denotes the sum of sampling and nonsampling error in the estimated probability.

Methodology for TSE Analysis (2)

B.   Obtain best estimates of each component error from sources with higher coverage and/or response rates

1) Vital Statistics data
2) American Community Survey (ACS)
3) National Health Interview Survey (NHIS) and NHIS-Provider Record Check Study (NHIS-PRC)

C. Generate a Monte Carlo simulated dataset using the best sources of component error to estimate total bias in vaccination coverage rate estimates and associated sampling error.

Three Different Weighted Estimates
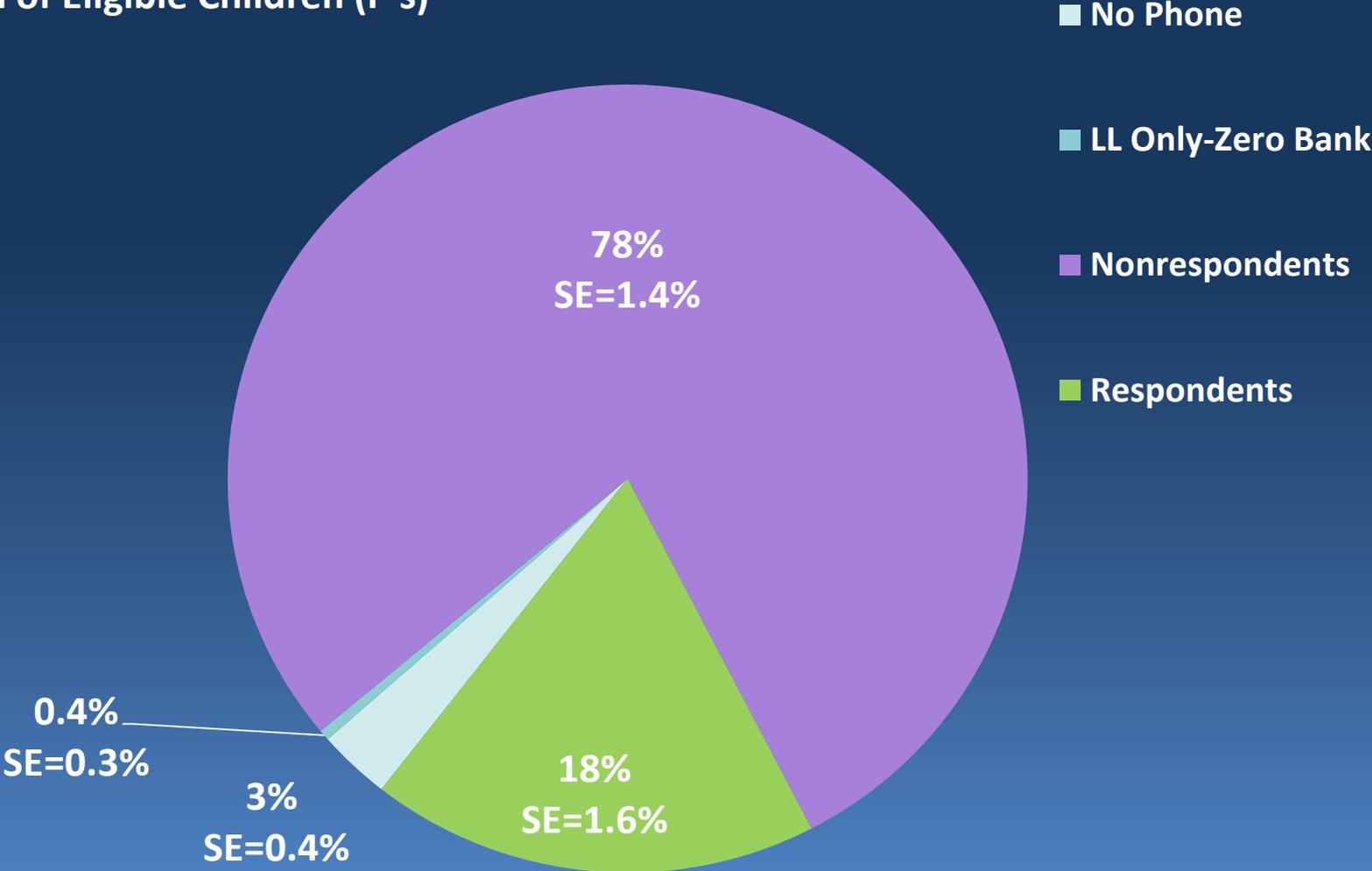
Design Weights
- Adjust for the selection probabilities
- Adjust for the multiplicity of telephone lines
- Adjust for the in household selection for the landline sample
- Adjust for the overlap of the landline and cell sample.

Unbiased Weights and **Shrinkage Weights**
- Full set of weighting adjustment including adjustments for selection probabilities, nonresponse, combining the landline and cell samples, and raking.
- Unbiased weights: used the true cell phone only cases to represent the cell phone only population
- Shrinkage weights: used the true cell phone only cases along with some borrowed landline cases to represent the cell phone only population
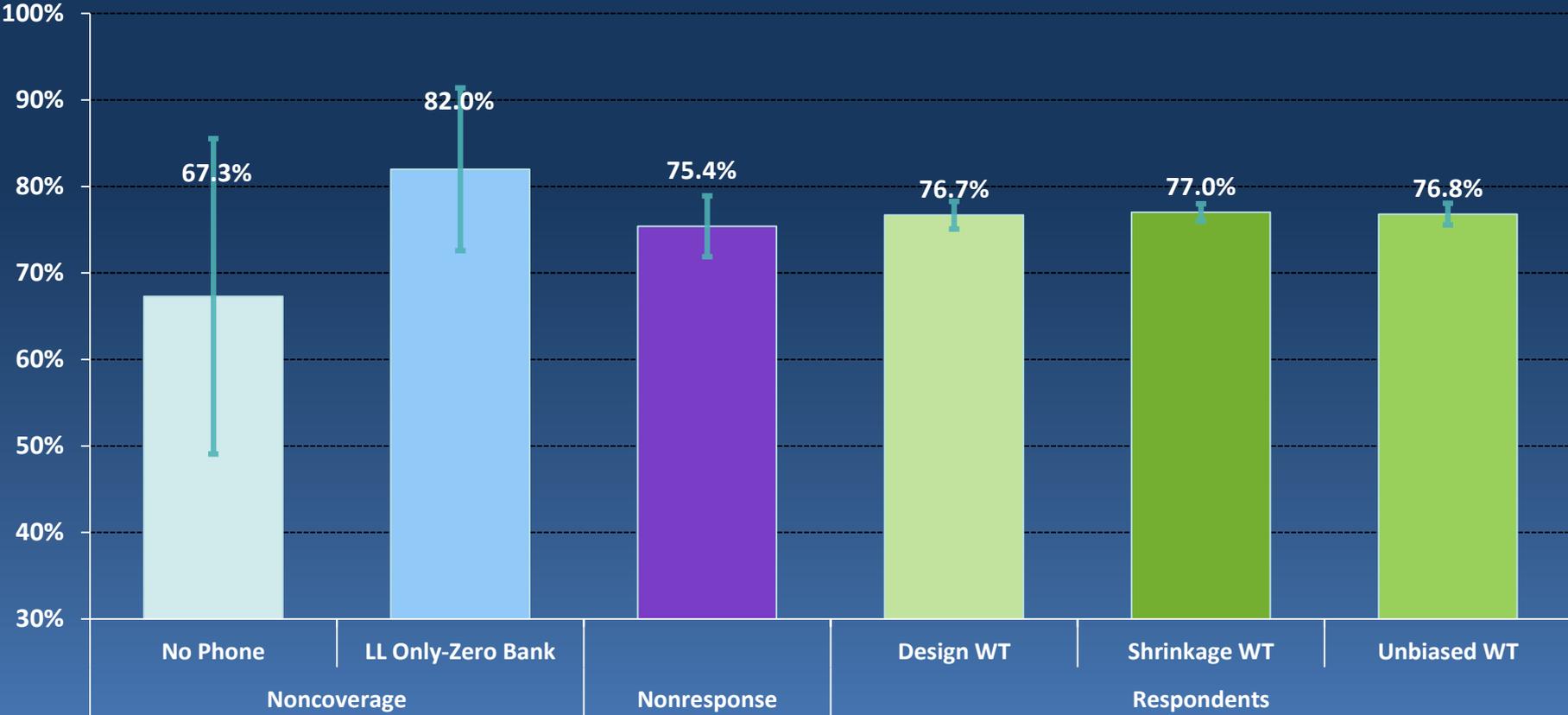
2011 Simulation Inputs -- NIS (1)

Distribution of Eligible Children (P's)

Legend:
- No Phone
- LL Only-Zero Bank
- Nonrespondents
- Respondents

78%
SE=1.4%

18%
SE=1.6%

3%
SE=0.4%

0.4%
SE=0.3%

2011 Simulation Inputs -- NIS (2)

Best Estimated Vaccination Coverage Rates for 4:3:1:3:3:1 Series (μ's)

No Phone	LL Only-Zero Bank		Design WT	Shrinkage WT	Unbiased WT
67.3%	82.0%	75.4%	76.7%	77.0%	76.8%
Noncoverage		Nonresponse	Respondents		

2011 Simulation Inputs -- NIS-Child (3)

Component stage	Phone status and weighting	4+ DTaP		1+ MMR		4:3:1:3:3:1	
		mean	SE	mean	SE	mean	SE
Non-Covered	No phone	67.3%	9.3%	82.0%	7.2%	67.3%	9.3%
	LL-only zero-bank	86.1%	4.3%	93.0%	3.0%	82.0%	4.8%
Non-Respondent	Non-APD	83.4%	1.6%	91.6%	1.2%	75.4%	1.8%
Respondent	NIS weighted (Design wt)	84.0%	0.7%	91.2%	0.5%	76.7%	0.8%
	NIS final weighted (Shrinkage wt)	84.6%	0.5%	91.6%	0.4%	77.0%	0.6%
	NIS final weighted (Unbiased wt)	84.4%	0.6%	91.5%	0.5%	76.8%	0.7%

Result 1: Estimated Total Bias Using Three Alternative Weights – 2011 NIS

Estimated Total Bias %

	4:3:1:3:3:1 Series	1+MMR	4+DTaP
Design WT	1.3	-0.1	0.9
Unbiased WT	1.4	0.3	1.3
Shrinkage WT	1.5	0.3	1.5

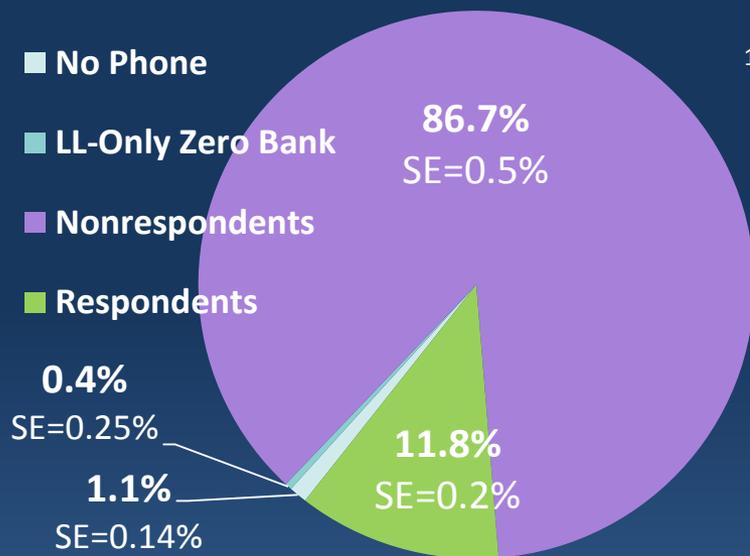■ Design WT ■ Unbiased WT ■ Shrinkage WT
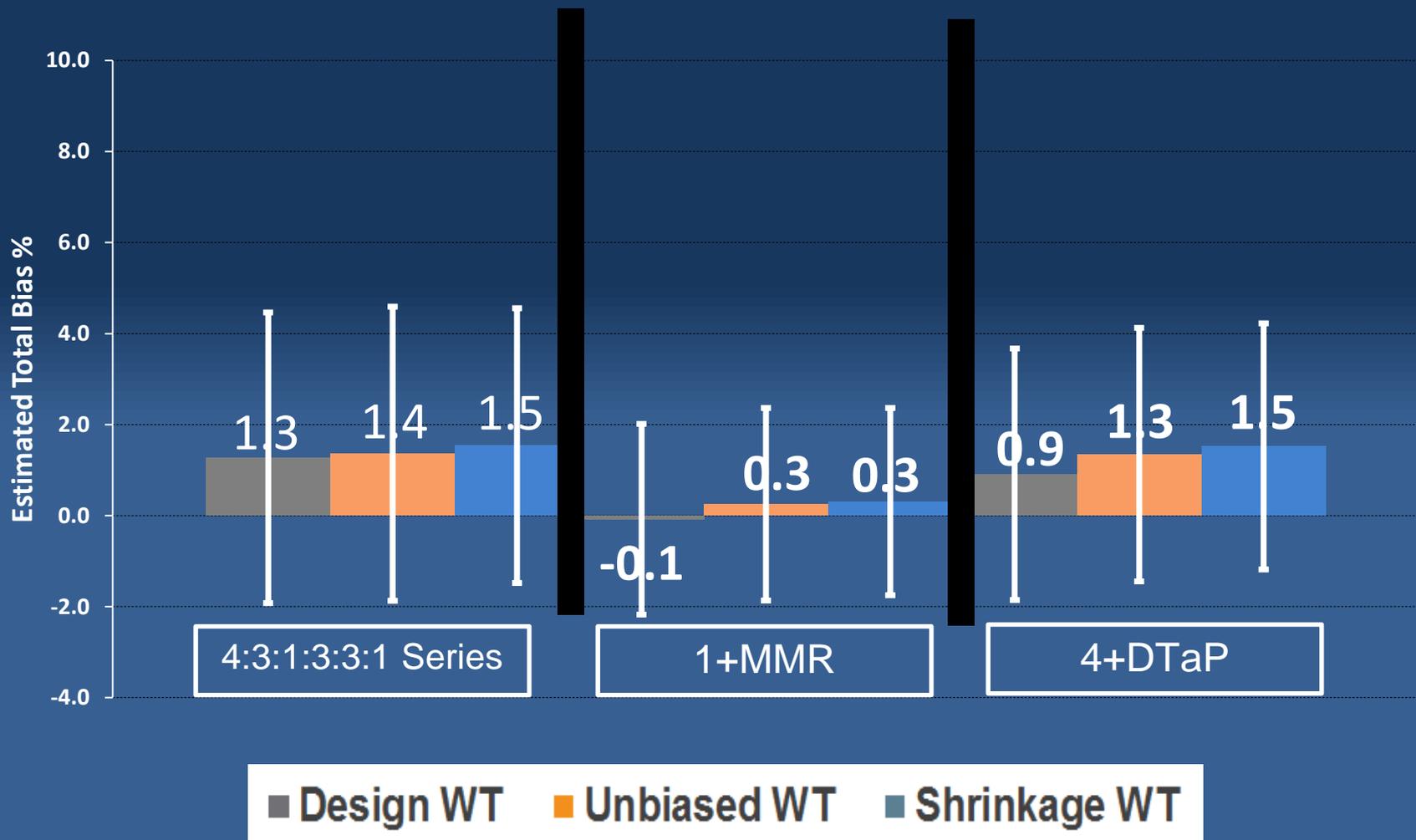
Result 2: Estimated Incremental Bias due to Noncoverage and Nonresponse – 2011 NIS

13
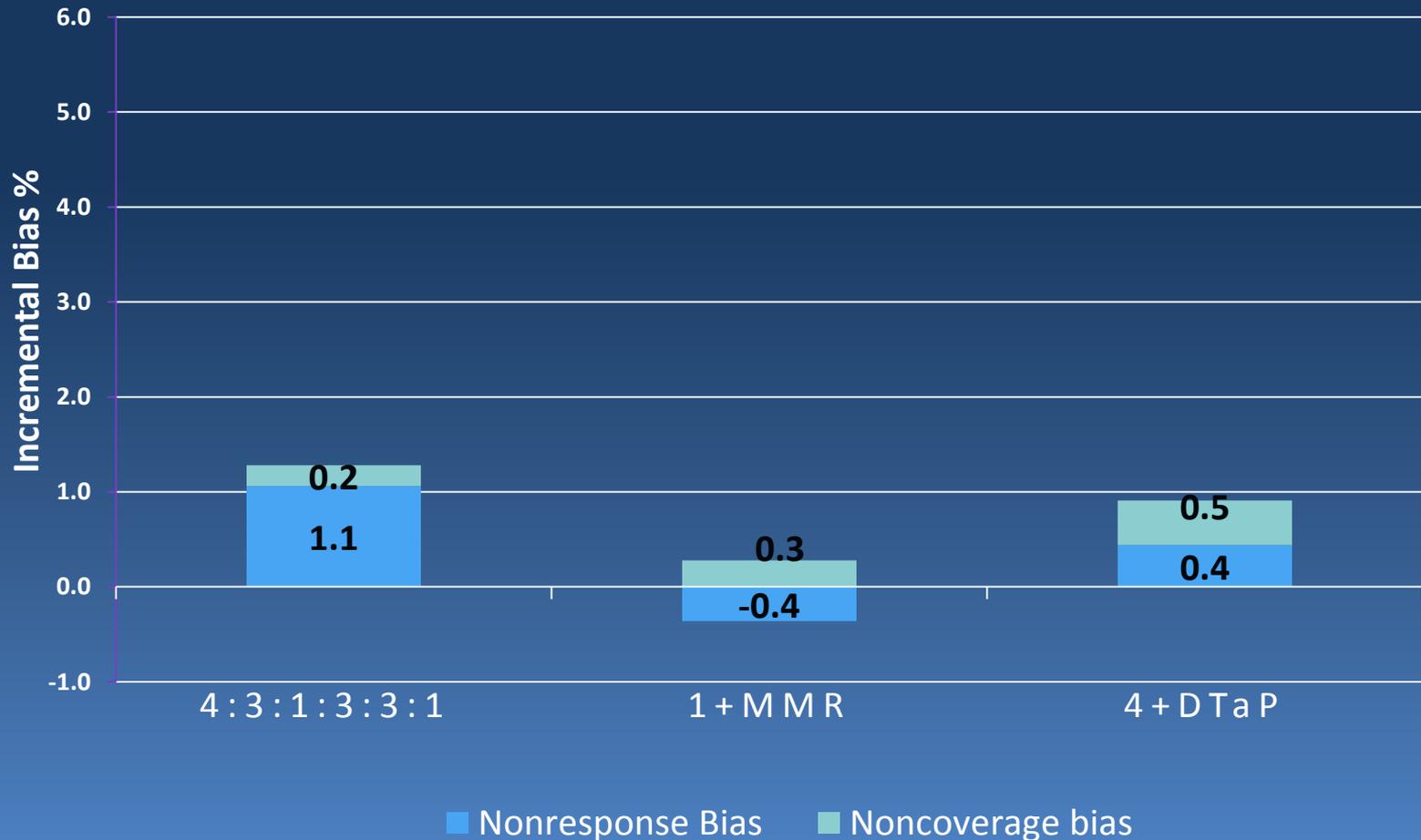
2011 Simulation Inputs -- NIS-Teen

Distribution of Eligible Teen (P's)

- No Phone
- LL-Only Zero Bank
- Nonrespondents
- Respondents

86.7% SE=0.5%

11.8% SE=0.2%

0.4% SE=0.25%

1.1% SE=0.14%

Best Estimated Vaccination Coverage Rates for 1+ Tdap (μ's)

No Phone: 52.72%	Zero Bank: 81.9%	Nonresponse: 74.9%	Design WT: 80.2% / Shrinkage WT: 78.2% / Unbiased WT: 78.0%
Noncoverage		Nonresponse	Respondents
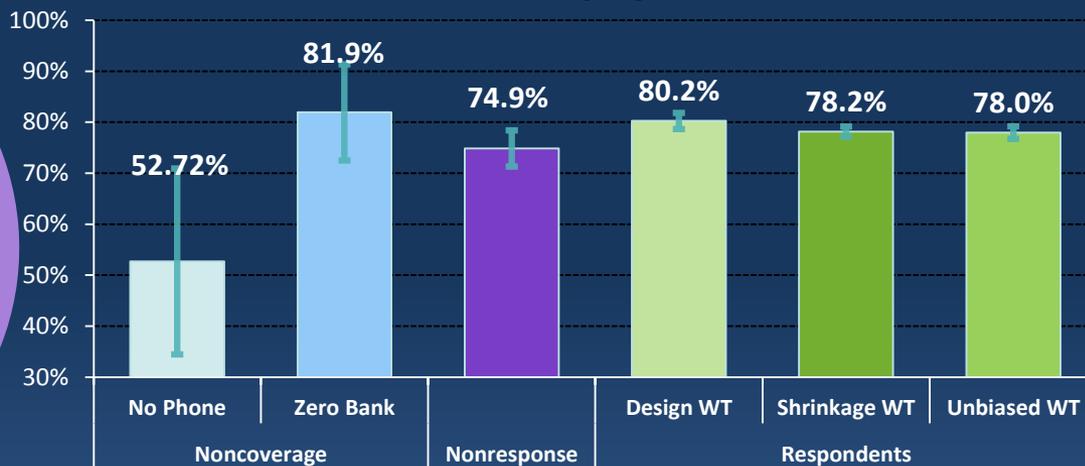
Component stage	Phone status and weighting	1+ Tdap		1+ MenACWY		1+ HPV among females	
		mean	SE	mean	SE	mean	SE
Non-Covered	No phone	52.7%	13.6%	54.0%	12.1%	67.0%	17.3%
	LL-only zero-bank	81.9%	3.6%	70.8%	4.3%	46.8%	6.3%
Non-Respondent	Non-APD	74.9%	1.1%	67.9%	1.2%	46.4%	1.8%
Respondent	NIS weighted (Design wt)	80.2%	0.5%	72.2%	0.6%	51.4%	1.0%
	NIS final weighted (Shrinkage wt)	78.2%	0.5%	70.5%	0.5%	53.0%	0.8%
	NIS final weighted (Unbiased wt)	78.0%	0.6%	69.9%	0.6%	52.8%	1.0%

Result 3: Estimated Total Bias using 3 Alternative Weights − 2011 NIS-Teen
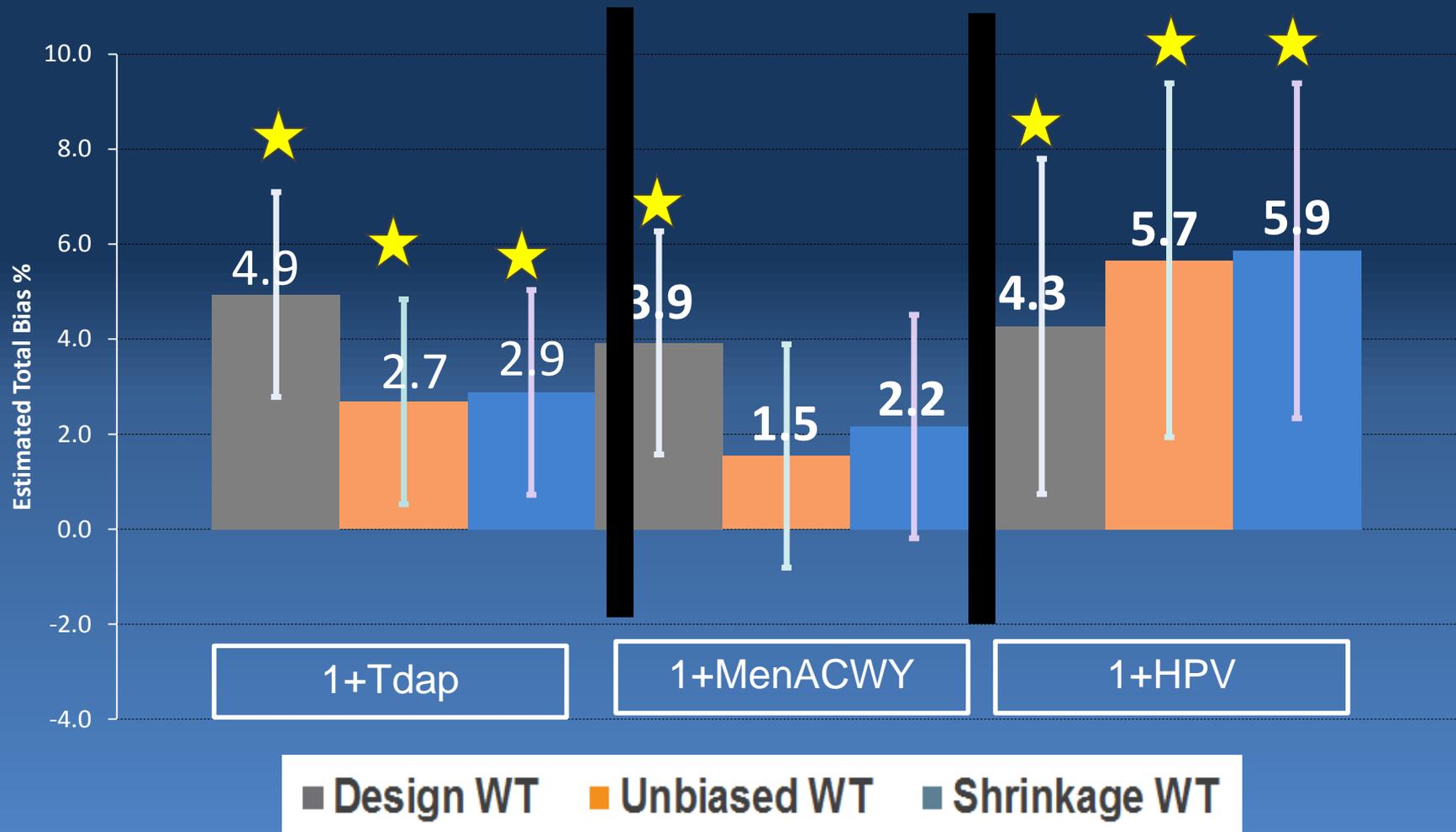
Statistically significant at the 95% confidence level

Result 4: Estimated Incremental Bias due to Noncoverage and Nonresponse – 2011 NIS-Teen
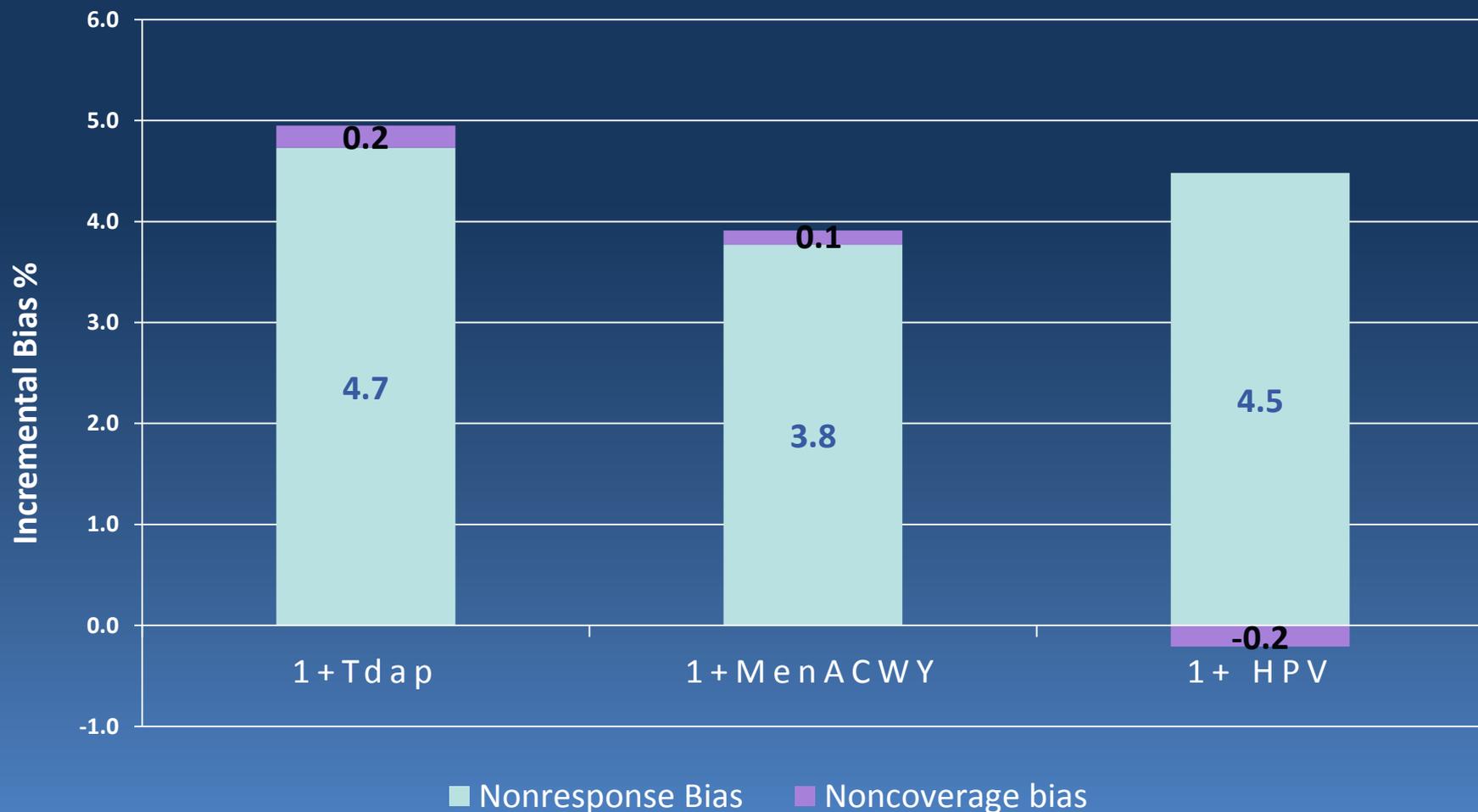
Discussion (1)

The estimated bias associated with population noncoverage is found to be small for both NIS-Child and NIS-Teen with the dual frame telephone sample design.

NIS-Child:

- The total bias in the 2011 vaccination coverage rates examined are quite small (<=1.5%) and not significantly different from zero, which is comparable to 2010 results.

NIS-Teen:

- Statistically significant biases were detected in the 2011 NIS-Teen vaccination coverage rates ranging from 2.9% to 5.9% (under shrinkage weights)
- The estimated bias from nonresponse dominate the total survey error.

Discussion (2)

- Use of shrinkage weights in 2011 was effective at reducing variance in vaccination coverage rate estimates without increasing bias.

- Limitation wise, as with other TSE studies, results are based on several assumptions, multiple sources of data, and models used in simulation.

- Total Survey Error analysis will be repeated using 2012 NIS Child and Teen data which has a larger cell-phone sample.

Thank You.

Meena Khare
> **e-mail: Mkhare@cdc.gov**

Vicki Pineau
> **e-mail: Pineau-Vicki@norc.org**