#### Total Survey Error Analysis: National Immunization Survey Adult COVID Module COVID-19 Vaccination Rates

#### May 11, 2022

Elizabeth Allen, NORC Vicki Pineau, NORC Kirk Wolter, NORC Jason Boim, NORC James A. Singleton, CDC Michael Chen, CDC David Yankey, CDC Jennifer Kriss, CDC Yi Mu, CDC

#### +NORC at the University of Chicago

\*The findings and conclusions in this presentation are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

 $\bigcirc$ 

 $\bigcirc$ 

### Agenda

- 01 Background and Overview
- 02 Coverage Error
- 03 Nonresponse Error
- 04 Reporting Error

05 Discussion

# Background and Overview



**Goal:** Investigate the potential sources for error within the National Immunization Survey-Adult COVID Module (NIS-ACM)

- The NIS is sponsored by the Centers for Disease Control and Prevention (CDC) and was implemented in 1994 to monitor vaccination coverage in children
  - NIS is an RDD cell-phone sample designed for reporting vaccination rates for states, D.C. Puerto Rico, USVI, Guam, and selected local areas
  - Provides national, state, and selected local area estimates of vaccination coverage rates
- NIS-ACM began in April-2021 to monitor COVID-19 vaccination status
  - Also asks about intent to vaccinate, vaccine hesitancy, and social behaviors
  - Key component of NIS-ACM is to understand vaccination disparity between different socio-demographic groups (e.g., race/ethnicity, poverty level), which is typically unavailable in the immunization counts from state records

**Goal:** Investigate the potential sources for error within the National Immunization Survey-Adult COVID Module (NIS-ACM)

- NIS-ACM uses the NIS sample as source for identifying eligible adults for NIS-ACM
  - NIS-ACM is fielded after all other modules within the NIS suite of surveys
  - Other modules are NIS-Child, NIS-Teen, NIS-CIM (Child Influenza Module), and CCM (Child COVID-19 Module)
- Weekly and monthly estimates are derived using a complex weighting design
  - Weighting steps include raking to population demographics as well as a final step that calibrates to Immunization Information System (IIS) vaccination count totals, specifically those who have had 1 or more COVID-19 vaccination doses
    - The calibration step is implemented for sex and age group, by state/jurisdiction
- Large sample pool allows for more stable estimates for subgroups of interest

 $\pm$ NORC

#### CDC Covid Data Tracker Vaccination Benchmark Data\*

- CDC estimates the number of people with at least one dose, fully vaccinated, and with a booster dose. CDC estimates are based on data that includes a dose number
- If a person received doses in more than one jurisdiction or at different providers within the same jurisdiction and given different person identifiers, CDC may not be able to link multiple unique person identifiers for different jurisdictions or providers to a single person
  - For example, the person's booster dose may appear to be their first dose when reported
- As of August 9, 2021, all entities have the ability to update or delete their previously submitted records.
- CDC has capped the percent of pop coverage metrics at 95%, which helps address potential
  overestimates of vaccination coverage due to 1<sup>st</sup>, 2<sup>nd</sup> and booster doses that were not linked
- Different reporting practices by jurisdiction can cause CDCs dose number estimates to differ from those reported by jurisdictions and federal entities

 $\pm$ NORC

#### Comparison of NIS-ACM estimates to benchmark

- Initial NIS-ACM national estimates were about 12 percentage points higher than the CDC COVID-19 tracker data
  - Differences have been reduced over time as the overall vaccination rate has increased



#### 1+ COVID-19 Vaccination Rates for Adults 18+: Early May – December 2021\*

\* Weighted Estimates using demographically post-stratified weights for each survey. Estimates reported in the presentation may diff somewhat from officially reported estimates for various reasons. For the ACM, estimates published by CDC also include a weighting calibration to IIS-based control totals.

#### Total Survey Error Framework

- Defined as the difference between the true, but unknown, population rate and the weighted survey estimate
- Sum of all errors that could be introduced: sampling and nonsampling errors
  - Sampling error: sample size, survey design, estimation
  - Nonsampling error: sampling-frame coverage, nonresponse, measurement
- This analysis focuses primarily on nonsampling errors, divided into three stages (sampling frame coverage, nonresponse, reporting error)

### Sampling-Frame Coverage Error



#### Sampling-frame coverage error

- NIS utilizes a cell-phone frame for selecting numbers to dial
  - Does not cover adults living in landline-only households or those who are phoneless
- Adults 18 and over live primarily in households with a cell-phone device
  - 97.3 percent of adults live in households with a cell-phone\*
    - 68.0 percent of adults live in households that only have a cell-phone,
    - 29.3 percent live in households with both a cell-phone and a landline
  - The phoneless population makes up only 0.5% of the adult population
- A majority of adults are covered with the cell-phone frame, but the difference in coverage varies by age group, race/ethnicity, and sex
  - Cell-phone frame has lower coverage rates of:
    - Older adults, especially 65 and older
    - Non-Hispanic Black
    - Females

**XNOR** 

#### Potential coverage error

- Most age groups have coverage over 95%
- 65+ age group has the lowest coverage in the cell-phone frame, of 89.1%
  - 9.2% of the 65+ age group live in landlineonly households

Age group	NHIS Wireless Estimate of Coverage
18-24	96.6
25-29	98.0
30-34	98.3
35-44	97.3
45-64	96.5
65+	89.1

#### coverage error

- Other demographics have more consistency
- Race/ethnicity shows very stable coverage of around 95%
  - Even though NH Black has lower coverage, still very high
- Females have slightly lower coverage than males, but minimally

Race/Ethnicity	NHIS Wireless Estimate of Coverage
Hispanic	95.0
White, non-Hispanic	95.8
Black, non-Hispanic	93.8
Asian, non-Hispanic	94.7
Other or multiple race, non-Hispanic	94.2
	NUIS Wirologg
Sex	Estimate of Coverage
Male	95.9
Female	94.7

### Nonresponse Error



#### Nonresponse error

- Multiple reasons someone may not respond to NIS-ACM
- Leverage saliency theory models the probability of someone responding to a survey
  - Survey topic: interest and/or saliency of topic to respondent
  - Sponsorship: who is funding the research
  - Incentive: monetary or nonmonetary
  - Survey mode: enjoyment/aversion to direct interaction with interviewer
- Those who are vaccinated or more at risk to get a severe case of COVID-19 may be more engaged in the topic, and more likely to respond
- Those who are unvaccinated and are not planning on getting vaccinated may be less likely to respond to the survey

 $\pm NORC$ 

- Age group comparison
  - Distribution of American Community Survey (ACS) versus design-weighted respondent distribution
    - Some small differences, but all within 4 percentage points
  - Larger differences associated with youngest and oldest age groups
  - Age group is used in raking for the final weights, and IIS calibration

Age Group	American Community Survey	Design-Weighted Distribution (Feb-2022)	Difference (ACM-ACS)
18-29	17.6%	21.0%	3.4%
30-39	15.6%	17.3%	1.7%
40-49	14.1%	15.9%	1.8%
50-64	27.8%	24.6%	-3.2%
65+	24.9%	21.2%	-3.7%

- Comparison of sex
  - Design-weighted distributions are very similar to ACS for female and male distributions
  - Sex is used in raking for the final weights, and IIS calibration

Sex	American Community Survey	Design-Weighted Distribution (Feb-2022)	Difference (ACM-ACS)
Female	51.4%	52.2%	0.8%
Male	48.6%	47.8%	-0.8%

- Race/ethnicity comparison
  - Larger differences associated with Hispanics, Non-Hispanic White
    - Hispanics appear to be underrepresented in the base-weighted distribution compared to ACS
    - NH Whites are slightly overrepresented
  - Race/ethnicity is used in raking for the final weights (not in the IIS calibration)

Race/Ethnicity	American Community Survey	Design-Weighted Distribution (Feb-2022)	Difference (ACM-ACS)
Hispanic	17.2%	12.0%	-5.2%
NH White	62.2%	65.0%	2.8%
NH Black	12.0%	12.3%	0.3%
NH Asian	6.0%	6.5%	0.5%
NH Other	2.7%	4.2%	1.5%

- Metropolitan Statistical Area (MSA) comparison
  - County defines MSA versus Non-MSA
  - Distribution of population based on MSA status is relatively similar
    - Slight underrepresentation of Non-MSA population in pool of respondents

MSA	American Community Survey	Design-Weighted Distribution (Feb-2022)	Difference (ACM-ACS)
MSA	85.2%	87.8%	2.6%
Non-MSA	14.8%	12.2%	-2.6%

- Social Vulnerability Index (SVI) comparison
  - Defined rankings by county, takes into account socio-demographics of each county
    - Those with high SVI rankings are more vulnerable populations, and have lower vaccination rates
  - Distribution of population in Low, Moderate, and High SVI ranking versus weighted respondent distribution are fairly similar

SVI Rank	American Community Survey	Design-Weighted Distribution (Feb-2022)	Difference (ACM-ACS)
Low	27.4%	27.9%	0.5%
Medium	38.6%	39.6%	1.0%
High	34.0%	32.5%	-1.5%

- Political Grouping comparison
  - Defined groups by county, not respondent
    - Based on the percent of vote for Biden or Trump in the November 2020 presidential election within the county, not the respondent's political affiliation
    - County groups were defined as Republican leaning, Democrat leaning, and Not Democrat or Republican Leaning
  - Distribution of population across the three groups is very similar to the distribution of weighted respondents

Political Grouping	American Community Survey	Design-Weighted Distribution (Feb-2022)	Difference (ACM-ACS)
Republican-Leaning Counties	31.0%	29.9%	-1.1%
Not Republican or Democrat Leaning	24.6%	24.5%	-0.1%
Democrat-Leaning Counties	44.4%	45.6%	1.2%

### Reporting Error



#### Reporting error

- Multiple reasons for measurement error
  - Respondent recall error
  - Social desirability
  - Satisficing, i.e., wanting to give the interviewer the "correct" answer
  - Question wording may be estimating a different outcome construct than intended
- Difficult to determine true over-/under-reporting without a direct record comparison, as is done with NIS-Child, NIS-Teen
  - Can look at other NIS modules for an indication of the potential for measurement error from respondent
  - NIS-Flu, which has parental reports of child influenza vaccinations, has shown an over-reporting estimate of approximately 4 percentage points

#### Reporting error

- Started looking at experiment of question wording.
  - Question wording may have some impact on COVID-19 vaccination reporting
  - Other surveys have looked at question wording and found a small decrease in the overall vaccination rates with a change in the question wording
  - Current wording of question is "HAVE YOU RECEIVED AT LEAST ONE DOSE OF A COVID-19 VACCINE?"
- Comparison of NIS-Teen provider record check results for COVID-19 vaccination for NIS-CCM
  - Analysis planned to compare the parental reports of COVID-19 vaccination versus healthcare provider records for teens aged 13-17 years
  - Teen COVID-19 vaccination rates are highly correlated with parental COVID-19 vaccination rates
  - Comparison of parental reported and provider records for the teens in the household will give insight into the potential measurement error that exists within the NIS-ACM
- CDC has funded additional research to assess adult self-reported vaccination
  - University of Colorado will conduct the analysis, comparing adult self-report with medical records of vaccination and IIS data

 $\pm NORC$ 

### Discussion and Next Steps



#### Discussion

- Each of the three stages could produce errors in the final estimates for NIS-ACM
- Sampling-frame coverage error
  - There are small differences in coverage, though greater for those who are 65+
- Response error
  - Nonresponse bias likely one driver for overestimation of the initial NIS-ACM estimates
  - Reasons for nonresponse may be topic saliency, current political environment (distrust in government agencies, known vaccination coverage differences by political party affiliation)
- Reporting error
  - Under-, Over-reporting could have an impact here

#### Next Steps

- Comparison of self-report to vaccination records
  - Initial assessment based on NIS-Teen provider record check for 2021, which will compare parental reports
    of COVID-19 vaccination with PRC results
  - Assessment of adult self-report vaccination status compared to medical records and IIS data, for the state
    of Colorado
- Potential vaccination question experiment will shed light on the impact of question wording
- Additional TSE analyses
  - Benchmarking to IIS data allows assessment of the potential for overall bias
  - Continue to review and assess possible sources of error and bias at each stage of the survey process as new external data sources and benchmarks become available

 $\pm$ NORC

### Questions?



A LANDMARK NORC STUDY SINCE 1972



## Thank you.

**Elizabeth Allen** Principal Statistician Allen-elizabeth@norc.org



