Ensuring Representativeness of a Sampled Subpopulation from a Probability Panel

Jared Coopersmith[1], Zachary Lewis[1], Randall K. Thomas[1], Mary Noorzai[1], Matthew Bramlett[2], James Dahlhamer[2], & Katherine Irimata[2]

[1] Ipsos Public Affairs

[2] National Center for Health Statistics

1

Study Background

Disclaimer: The Rapid Surveys System (RSS) is conducted by the Centers for Disease Control and Prevention's (CDC's) National Center for Health Statistics (NCHS), under contract GS00F123DA with Ipsos with funding from CDC/NCHS and other cosponsors. The findings and conclusions in this presentation are those of the authors and do not necessarily represent the official position of CDC/NCHS or other funding partners.

Study Background

The Rapid Surveys System (RSS) established by the National Center for Health Statistics (NCHS) collects survey data from two probability-based online survey panels—Ipsos' KnowledgePanel (KP) and NORC's AmeriSpeak panel—to generate national population estimates of health-related experiences and behaviors.

In developing the RSS, NCHS was interested in more rapid assessment of emergent topics than what was possible with their existing large-scale data collection programs.

Study Background

RSS Rounds 1 – 4: Aimed for representativity of the U.S. adult population aged 18 and older

- Both Ipsos' KnowledgePanel and NORC's AmeriSpeak probability-based panels are designed to represent adults aged 18+
    - Standard sampling and weighting procedures

RSS Round 5: Intended to represent the population of U.S. children aged 0-17, with information reported by a parent or guardian

- Necessitated a non-standard sampling and weighting approach

KP Background and Study Overview

Background on KP

- **Built primarily using random sampling of addresses from an address-based sampling (ABS) frame (small Latino subset also uses RDD)**

- **Participants are provided with small incentives to participate in surveys**

- **Those without internet access are provided with a tablet with limited internet access to complete surveys online**

Ipsos' KP approach to producing a nationally representative sample of children with total survey error minimized combined the following:

- **A 'targeted' sample of parents of children aged 0-17, identified through panel profile data**

- **A general 'non-targeted' sample from the remainder of the panel**

Study Background

RSS-5 required interviews with various types of parental roles

- Parents (including stepparents and foster parents)

- Legal guardians (grandparents, aunts/uncles, etc.)

Population benchmarks were absent or inconsistent with this definition of the parental role (e.g., the National Health Interview Survey only captures 'parents')

- We examine how supplementing with benchmarks for the general adult population from U.S. Census Bureau surveys can be used to estimate the target population

Method

Method – RSS-5 Sample Design

Goal: Produce a nationally representative sample of approximately 4,000 children aged 0-17 by first selecting a nationally representative sample of parents or guardians who will respond on behalf of their children

We had multiple stages in this process:

1. **Selected a large general population sample (Sample 1) of 24,000 members of the KnowledgePanel**

    - **Used Probability Proportional to Size (PPS) sample selection technique**

    - **Results in a demographically-balanced sample representative of all U.S. households regardless of the presence of children**

Method – RSS-5 Sample Design

2. Split this large, demographically-balanced sample into two sub-groups

- 'Targeted' subgroup: All households who indicated having at least one child in the household based on profile data (N = 6,750)

- 'Non-targeted' subgroup: All households whose profile data did not indicate the presence of children (N = 17,250)

  - Included those reporting no children and those who failed to include this information in their profile data

Method – RSS-5 Fielding

3. **Sub-selection and survey assignment**

- **Targeted subgroup: 6,568 (of the 6,750 selected) were assigned to receive the online survey**

  - **182 were unavailable to be assigned after selection (some attrition, some away on vacation)**

- **Non-targeted subgroup: 10,000 were randomly selected for fielding from the 17,250 total selected**

  - **9,901 were actually assigned to the survey (99 unavailable)**

  - **We assumed a very low qualification rate in the non-targeted group, hence more efficient to field less than the full 17,250**

Method – RSS-5 Fielding

Completes: 3,935 qualified cleaned, completed cases – 3,681 (93.5%) respondents took the English version of the survey, and 254 (6.5%) took the Spanish version.

Field Period:

- **Soft launch:**   **July 8, 2024**

- **Full scale fielding:**   **July 11, 2024 – July 31, 2024**

- **Survey mode:**   **Online web-based survey**

- **Median interview length:**   **19.4 minutes**

Method – RSS-5 Fielding

The table below summarizes the numbers of sample members selected, assigned, completing the screener section, and the final clean, qualified completes. Sample 2 was selected but only the Targeted sub-group was assigned to meet the qualified completes goal.

		Selected	Assigned	Screener Completes	Cleaned Qualified Completes
Sample 1	Targeted	6,750	6,568	4,217	3,170
	Non-targeted	17,250	9,901	7,597	231
	Total	24,000	16,469	11,814	3,401
Sample 2	Targeted	1,236	1,235	723	534
	Non-targeted	4,764	0	0	0
	Total	6,000	1,235	723	534
Overall Total		30,000	17,704	12,537	3,935

Method – RSS-5 Questionnaire Topics

Round 5 – The parents/guardians of children aged 0 to 17 were asked questions about a single child in the household (randomly selected child if more than one child in HH). The questions regarding the child included:

- **Disabilities**

- **Healthcare visits and experiences**

- **Current medications**

- **Stressors and support (physical, emotional, educational, social)**

- **Vaccination experience**

Results

Results 1 – Why did some targeted sample disqualify?

941 (19%) of the 4,940 targeted sample screener completes disqualified

- **367 (7%) disqualified because they reported no 0-17-year-old children in the household**

- **574 (12%) disqualified because the panelist was not a parent/legal guardian of any children 0-17 in the household**

    - **47% were grandparents**

    - **31% were siblings**

    - **10% were aunts or uncles**

    - **Remaining 12% were other family/non-family members**

Results 2 – Was targeting related to HH characteristics?

Parents/guardians from non-targeted sample HHs had a lower income and were more likely to be the only adult in the HH ($p < .001$). The non-targeted sample also had a somewhat higher proportion of having an older youngest child (aged 12-17; $p < .07$).

Sample 1 – Unweighted HH Characteristics
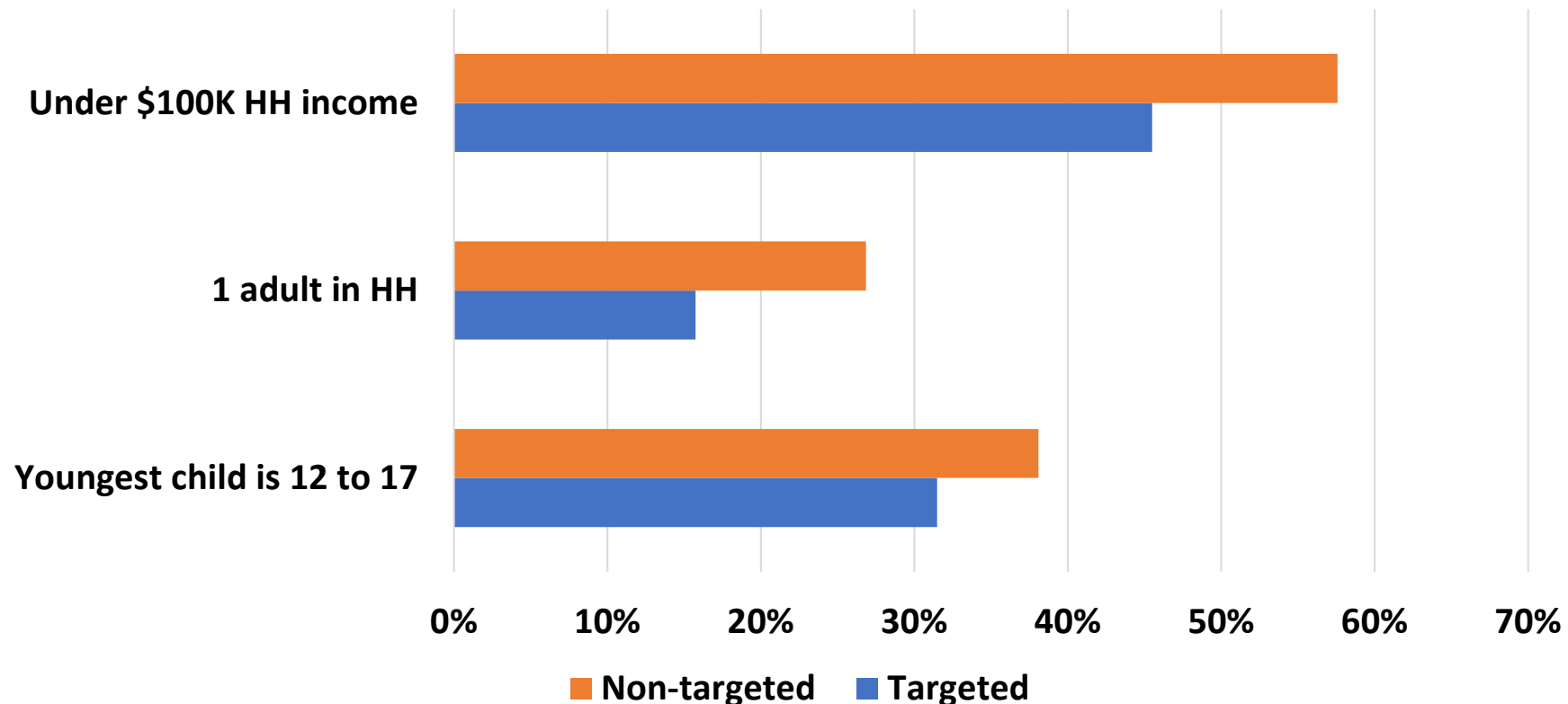
Results 3 – Was targeting related to parent characteristics?

Parents/guardians from targeted sample were more likely to be married, non-Hispanic White, aged 30 to 44, have a Bachelor's degree or higher, and be in a more traditional parental role (vs. another relationship, such as grandparent; *p* < .005).

Sample 1 – Unweighted Parent/Guardian Characteristics
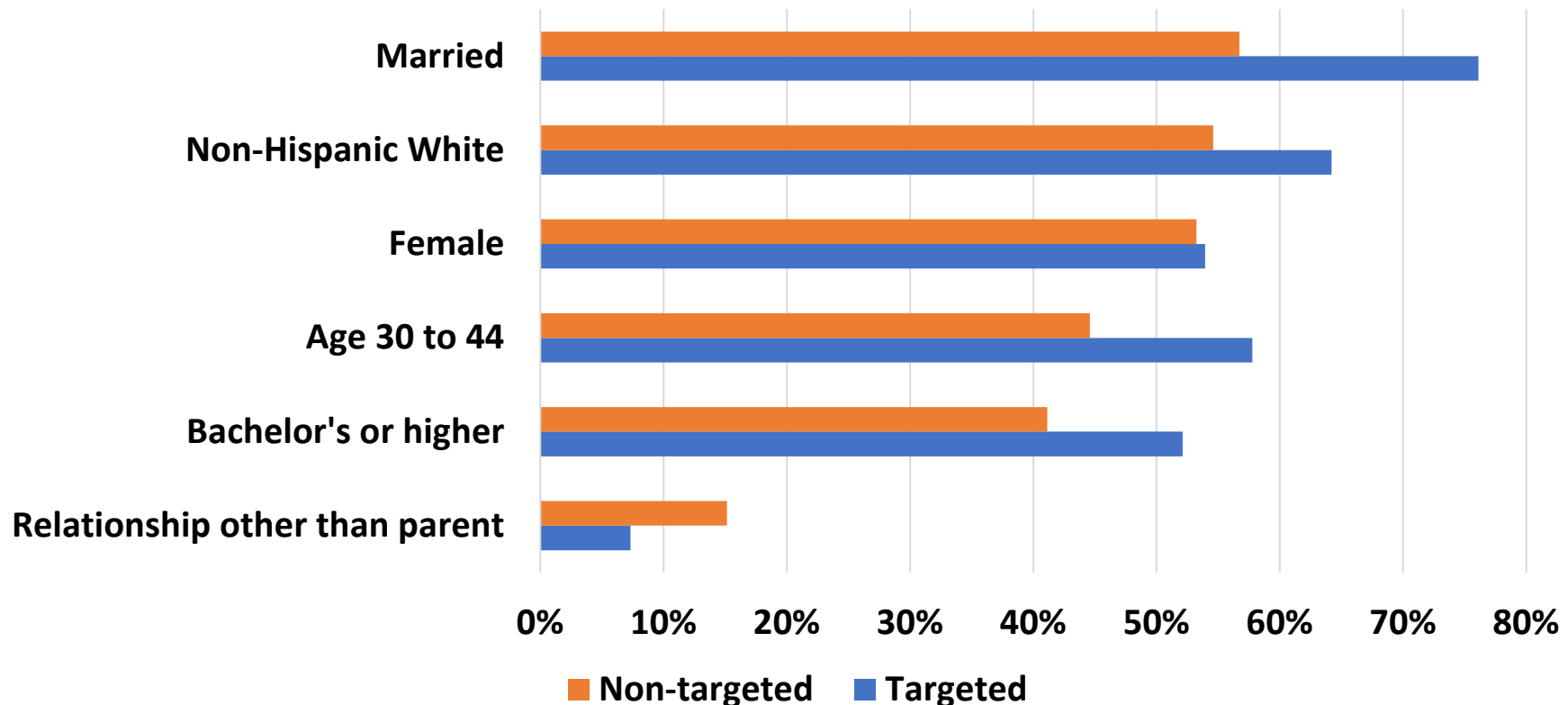
# Results 4 – Was targeting related to child health?

Despite differences in HH and parent characteristics, we found no significant differences in child health-related variables. Asthma had a marginally significant difference, with those in the targeted subgroup having a lower prevalence of asthma ($p < .07$).

Sample 1 – Unweighted Child Health
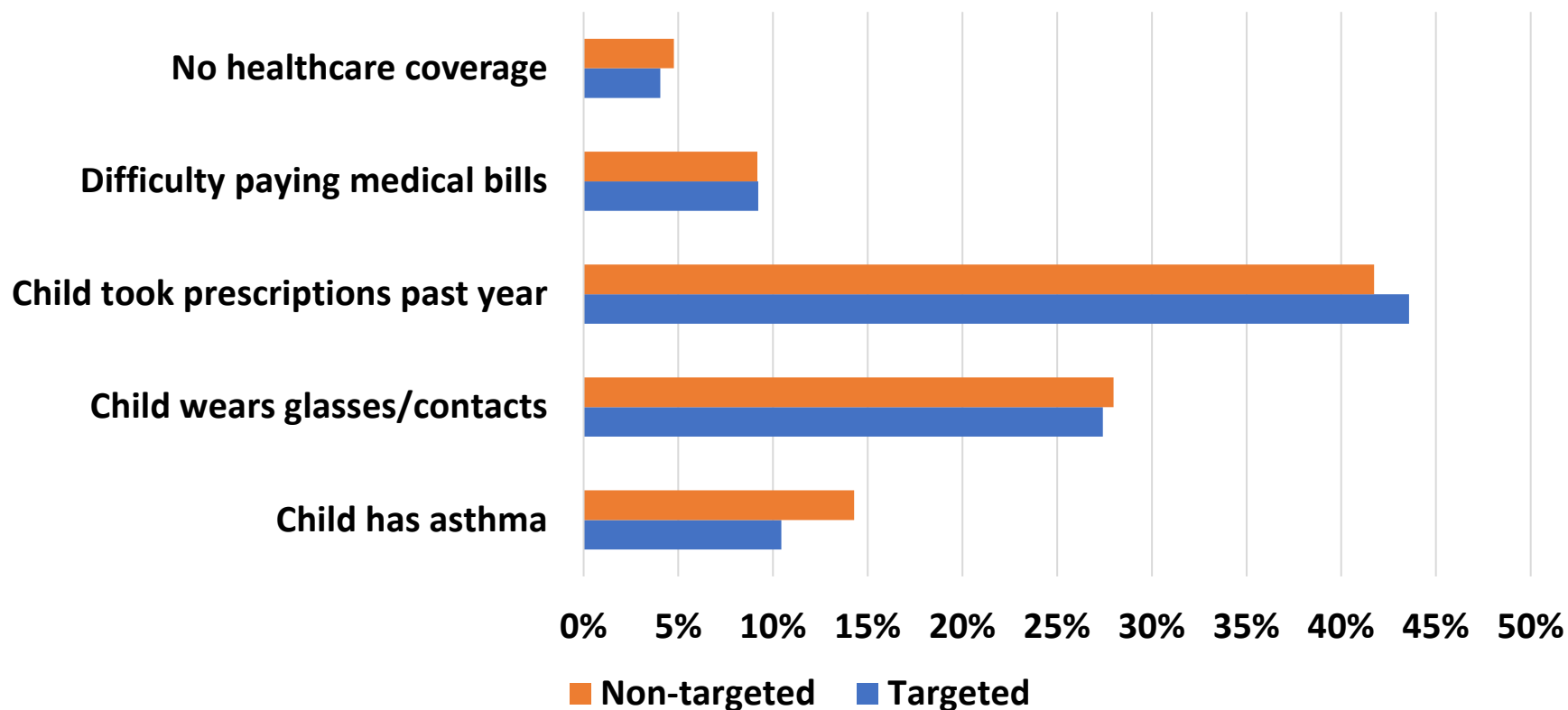
Results 5 – How did results compare to NHIS benchmarks?
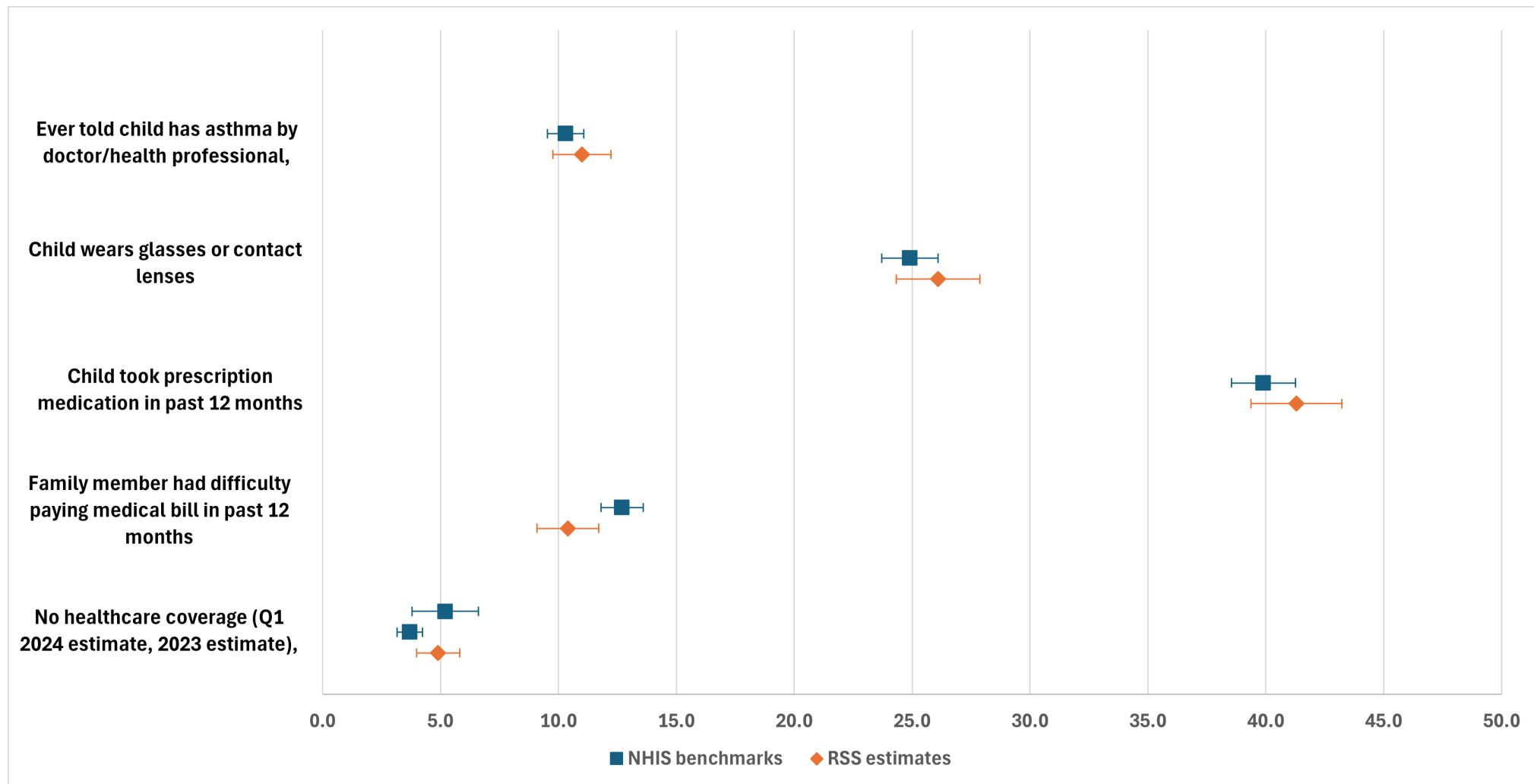
Note: 2023 NHIS estimates produced using SAS survey procedures controlling for PSUs and sampling strata. Q1 2024 healthcare coverage estimate source: "Health Insurance Coverage: Early Release of Quarterly Estimates From the National Health Interview Survey, January 2023–March 2024" https://www.cdc.gov/nchs/data/nhis/earlyrelease/Quarterly_Estimates_2024_Q11.pdf. NHIS data is from 2023 or Q1 2024, so some differences in estimates could result from time effects rather than sample designs. Full-year 2024 NHIS data is not yet available.

Discussion

Conclusions and Discussion

- **Qualification rate in the targeted group was lower than predicted**

  - **81% of screener completes qualified for the study (~90% was expected)**

- **Some significant differences between targeted versus non-targeted sub-groups in household characteristics**

- **Significant differences between targeted versus non-targeted sub-groups in parent/guardian characteristics**

- **Few differences between targeted versus non-targeted sub-groups in child health-related issues**

- **Strong correspondence between NHIS results with RSS-5 results**

Thank you!

Jared Coopersmith

Jared.Coopersmith@Ipsos.com

Ipsos U.S. Public Affairs is a Charter Member of AAPOR's Transparency Initiative