Weighting with the Redesigned National Health Interview Survey

Federal Committee on Statistical Methodology Research and Policy Conference
September 21, 2020

James M. Dahlhamer, Matthew D. Bramlett, & Jonaki Bose
National Center for Health Statistics

Ronaldo Iachan & Matt Jans
ICF
National Health Interview Survey (NHIS)

- **Agency**: Department of Health and Human Services (DHHS); Centers for Disease Control and Prevention (CDC); National Center for Health Statistics

- **Purpose**: To monitor the health of the US population through the collection and analysis of data on a broad range of health topics

- **Sample**: Complex sample of the civilian noninstitutionalized US population

- **Mode**: In-person interviews -- follow-up by telephone if needed

- **Data collection**: Continuous by Census field representatives

- **Target sample size**: 27,000 sample adults & 9,000 sample children

- **Content redesign start**: January 2019
2019 NHIS Questionnaire Redesign

- **Goals of the redesign**
  - Reduce the overall length of the interview
  - Improve the relevancy of covered health topics
  - Focus on leading causes of morbidity and mortality, targets of health promotion initiatives, and risk and protective factors
  - Harmonize overlapping content with other federal health surveys
  - Reduce respondent burden

- **Eliminated family interview; now a short household roster followed by selection of sample adult (18+) and sample child (<18)**
  - Sample adult and sample child are now only people we collect detailed health information on
Time for a New Weighting Approach?

- Questionnaire redesign provided opportunity to reconsider weighting approach, especially how we adjust for nonresponse (NR)
  - Not updated in previous ~20 years
  - Declines in NHIS household response rates (~90% to ~65%)
  - Likely break some trends due to content changes

- Other developments
  - Causes of survey NR have likely changed over time
  - Statistical methods and computing power have improved
  - Increasing availability of auxiliary data
Simplified Look at *Past* Weighting Approach

- **Final household base weight**
- **Household NR adjustment** *(segment-level adjustment)*
- **Post-stratification to population control totals based on age, sex, race/ethnicity**
- **Final person weight**

**Adult probability of selection adjustment**
- **Adult level NR adjustment** *(segment-level adjustment)*
- **Interim sample adult weight**
- **Post-stratification** to population control totals based on age, sex, race/ethnicity
- **Final sample adult weight**

**Child probability of selection adjustment**
- **Child level NR adjustment** *(segment-level adjustment)*
- **Interim sample child weight**
- **Post-stratification** to population control totals based on age, sex, race/ethnicity
- **Final sample child weight**
ICF Contract

- Explored various models of survey response incorporating contextual data (county-, Census tract-level measures) and paradata related to both response and key health measures:
  - Logistic regression
  - *Multilevel logistic regression*
  - Random forest prediction
  - Least absolute shrinkage & selection operator (LASSO) machine learning

- Explored *raking* versus traditional post-stratification
  - Old post stratification variables: age, sex, race/ethnicity
  - Raking: age, sex, race/ethnicity plus:
    - Education, employment status, MSA status and/or Census division
Simplified Look at *New* Weighting Approach

- Final household base weight
- Household NR adjustment *(based on multi-level logistic regression)*
  - Adult probability of selection adjustment
  - Adult level NR adjustment *(based on multi-level logistic regression)*
  - Interim sample adult weight
  - Raking *(age, sex, race/ethnicity, education, MSA status, Census Division)*
  - Final sample adult weight

- Child probability of selection adjustment
- Child level NR adjustment *(based on multi-level logistic regression)*
  - Child level NR adjustment *(based on multi-level logistic regression)*
  - Interim sample child weight
  - Raking *(age, sex, race/ethnicity, MSA status, Census Division)*
  - Final sample child weight
Modeling Survey Response

FCSM 2020
Screening Variables for Inclusion in Models of Response (1)

- Step 1: Explore associations between auxiliary variables and response
  - Response defined as response (1) vs. nonresponse (0)
  - Variables significantly associated with response (p < .10) move to the next step
  - This process performed for household, sample adult, and sample child response

- Step 2: Explore associations between auxiliary variables (screened in from step 1) and a set of key health indicators (KHIs)
  - Variables significantly associated with KHIs (p < .10) in at least 3 of 5 health domains for the adult and 2 of 3 health domains for the child move to the modeling stage
Auxiliary Variables: Data Sources

- **Contact History Instrument (CHI)**
  - Data collected on every contact attempt
  - Variables summarized to household/respondent level
  - Example: Whether or not householder(s) expressed time constraints

- **Neighborhood Observation Instrument (NOI)**
  - Data collected from first observation of sample unit
  - Example: Does the sample unit have any indication that the residents are smokers?

- **Census Planning Database**
  - Decennial Census 2010, American Community Survey 5-year estimates (2013-17)
  - Census tract-level measures
  - Example: Percentage of ACS population that is 65 years old or over

- **Area Health Resource File (AHRF)**
  - Contains more than 6,000 variables related to health care access
  - County-level measures
  - Example: Number of medical doctors per 100,000 county residents
Adult Key Health Indicators

**Health care service use**
- ER visit, past year
- Doctor visit, past year
- Flu vaccination, past year
- Mental health counseling/therapy, past year

**Health care access**
- Usual source of care
- Skipped doses of prescription meds to save money

**Health behaviors**
- Current smoker
- Obese

**Health insurance coverage**
- Public coverage
- Private Coverage
- Uninsured

**Health status**
- Functional disability
- Asthma episode, past year
- Hypertension
- Excellent/very good health
Child Key Health Indicators

**Health care service use**
- ER visit, past year
- Doctor visit, past year
- Flu vaccination, past year
- Mental health counseling/therapy, past year

**Health status**
- Disability
- Asthma episode, past year
- Excellent/very good health
- Current ADD/ADHD
- Stressful life events indicator

**Health care access**
- Usual source of care
- Public coverage
- Private Coverage
- Uninsured
Screening Variables for Inclusion in Models of Response (2)

- **Step 1:** Explore associations between auxiliary variables and response
  - Response defined as response (1) vs. nonresponse (0)
  - Variables significantly associated with response ($p < .10$) move to the next step
  - This process performed for household, sample adult, and sample child response

- **Step 2:** Explore associations between auxiliary variables (screened in from step 1) and a set of key health indicators (KHIs)
  - Variables significantly associated with KHIs ($p < .10$) in at least 3 of 5 health domains for the adult and 2 of 3 health domains for the child move to the modeling stage

- ~60 of the 213 variables explored made it to the modeling stage
  - Household (n=48)
  - Adult (n=48)
  - Child (n=26)
Modeling Survey Response

- Step 1: Estimate a standard binomial logistic regression of response (response vs. nonresponse)
  - Includes variables retained from prior screening steps
  - Backward elimination ($p < .10$)
    - Variables retained in this procedure are carried forward to a multi-level logistic regression of response

- Step 2: Estimate a multi-level logistic regression of response (response vs. nonresponse)
  - Random effects for Census tract
  - Fixed effects for variables retained from Step 1
  - Output predicted probabilities (response propensities)
Forming and Applying Nonresponse Adjustments

FCSM 2020
Household Nonresponse Adjustment

- Step 1: Order the response propensities from low to high and group them into quintiles

- Step 2: Form the NR adjustment factor by taking the inverse of the median response propensity within each propensity quintile: $1 / \text{median RP}$

- Step 3: Multiply the final household base weight by the NR adjustment factor
Nonresponse Adjustment Factors: Household, 2019

<table>
<thead>
<tr>
<th>Response Propensity Quintiles</th>
<th>1 (low)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5 (high)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonresponse adjustment factor formed by taking the inverse of the median response propensity within each quintile.</td>
<td>3.44</td>
<td>1.92</td>
<td>1.40</td>
<td>1.23</td>
<td>1.11</td>
</tr>
</tbody>
</table>
Sample Adult and Sample Child Nonresponse Adjustments

- The NR-adjusted household weight is the starting point for the sample adult and sample child weight

- Step 1: Multiply the NR-adjusted household weight by inverse of adult (child) probability of selection
  - This becomes the adult (child) base weight

- Step 2: Group response propensities from the final adult (child) response model into quintiles

- Step 3: Form the NR adjustment factor by taking the inverse of the median response propensity within each propensity quintile: 1 / median RP

- Step 4: Multiply the adult (child) base weight by the NR adjustment factor
## Nonresponse Adjustment Factors: Sample Adult and Sample Child, 2019

<table>
<thead>
<tr>
<th></th>
<th>Response Propensity Quintiles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 (low)</td>
</tr>
<tr>
<td><strong>Sample adult</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.25</td>
</tr>
<tr>
<td><strong>Sample child</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.24</td>
</tr>
</tbody>
</table>

Nonresponse adjustment factor formed by taking the inverse of the median response propensity within each quintile.
Raking

FCSM 2020
Raking (1)

- Adjusting sample weights so that marginal totals of the adjusted weights on specified characteristics (e.g., age, sex, race/ethnicity) agree with corresponding totals for the population

- Most often used to reduce biases from nonresponse and noncoverage

- Iterative process by which all variables are considered in turn. For each variable, weights are adjusted to align the survey marginals with population marginals
Raking (2)

- More flexible than post-stratification (PS)
  - PS requires control totals for ALL cells of a cross-classification
    - PS can spread the sample too thinly over large number of adjustment cells
- Thus, raking can typically handle more variables than PS
Raking Variables

- Various combinations of the following variables were explored: age, sex, race/ethnicity, employment status (adult only), education (adult only), MSA status, and Census division

  • Final sample adult raking dimensions:
    - Age by sex (18 cells)
    - Age by race and ethnicity (26 cells)
    - Education (4 cells)
    - MSA status by Census division (23 cells)

  • Final sample child raking dimensions:
    - Age by sex (10 cells)
    - Age by race and ethnicity (15 cells)
    - MSA status (3 cells)
    - Census division (9 cells)
Final Weights

FCSM 2020
Sample Adult Weight: 2019

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Min.</th>
<th>1st Quartile</th>
<th>Median</th>
<th>Mean</th>
<th>3rd Quartile</th>
<th>Max.</th>
<th>CV (%)</th>
<th>DEFF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Old approach</td>
<td>31,997</td>
<td>644</td>
<td>4,214</td>
<td>6,634</td>
<td>7,842</td>
<td>9,815</td>
<td>121,990</td>
<td>68</td>
<td>1.46</td>
</tr>
<tr>
<td>New approach</td>
<td>31,997</td>
<td>629</td>
<td>3,958</td>
<td>6,254</td>
<td>7,842</td>
<td>9,351</td>
<td>154,456</td>
<td>80</td>
<td>1.64</td>
</tr>
</tbody>
</table>

CV = coefficient of variation
DEFF = design effect
To constrain variance, we explored capping the household NR adjustment at:

- 3.0
- 2.5
- 2.0

Trade-off: While capping will reduce the variance of the weights, it will likely result in an increase in NR bias.
# Capping Nonresponse Adjustment Factors: Household, 2019

<table>
<thead>
<tr>
<th></th>
<th>Response Propensity Quintiles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 (low)</td>
</tr>
<tr>
<td>Uncapped</td>
<td>3.44</td>
</tr>
<tr>
<td>Cap at 3.0</td>
<td>3.00</td>
</tr>
<tr>
<td>Cap at 2.5</td>
<td>2.50</td>
</tr>
<tr>
<td>Cap at 2.0</td>
<td>2.00</td>
</tr>
</tbody>
</table>

Nonresponse adjustment factor formed by taking the inverse of the median response propensity within each quintile.
Sample Adult Weight: 2019

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Min.</th>
<th>1st Quartile</th>
<th>Median</th>
<th>Mean</th>
<th>3rd Quartile</th>
<th>Max.</th>
<th>CV (%)</th>
<th>DEFF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Past approach</td>
<td>31,997</td>
<td>644</td>
<td>4,214</td>
<td>6,634</td>
<td>7,842</td>
<td>9,815</td>
<td>121,990</td>
<td>68</td>
<td>1.46</td>
</tr>
<tr>
<td>New approach (uncapped)</td>
<td>31,997</td>
<td>629</td>
<td>3,958</td>
<td>6,254</td>
<td>7,842</td>
<td>9,351</td>
<td>154,456</td>
<td>80</td>
<td>1.64</td>
</tr>
<tr>
<td>New, cap 3.0</td>
<td>31,997</td>
<td>640</td>
<td>4,027</td>
<td>6,357</td>
<td>7,842</td>
<td>9,455</td>
<td>139,764</td>
<td>76</td>
<td>1.58</td>
</tr>
<tr>
<td>New, cap 2.5*</td>
<td>31,997</td>
<td>652</td>
<td>4,107</td>
<td>6,460</td>
<td>7,842</td>
<td>9,551</td>
<td>121,705</td>
<td>73</td>
<td>1.53</td>
</tr>
<tr>
<td>New, cap 2.0</td>
<td>31,997</td>
<td>665</td>
<td>4,189</td>
<td>6,527</td>
<td>7,842</td>
<td>9,657</td>
<td>102,032</td>
<td>70</td>
<td>1.49</td>
</tr>
</tbody>
</table>

CV = coefficient of variation
DEFF = design effect

*Weight included on 2019 sample adult datafile
How Did We Decide?

- Relative to the uncapped weight, looked at mean squared error (MSE) for 15 adult KHIs (and 13 child KHIs) across the different capping levels

- For all 3 cap levels we saw a reduction in MSE relative to the uncapped weight
  - Any increases in bias were outpaced by reductions in variance

- Overall, capping at 2.5 provided the best bias-variance tradeoff relative to the uncapped weight
Impact on Estimates

FCSM 2020
## Comparison of Select Adult Estimates Using the Old and New Sample Adult Weight: 2019

| Outcome                        | Base Weight | Old Weight | New Weight | |New – Old| |
|--------------------------------|-------------|------------|------------|-----------------|----------|
|                                | %           | %          | SE         | %               | SE       | |
| Uninsured                      | 8.6         | 10.2       | 0.24       | 11.0            | 0.27     | 0.8     |
| Doctor visit, past year        | 86.7        | 85.3       | 0.26       | 84.9            | 0.27     | 0.4     |
| Flu vaccination                | 50.6        | 47.7       | 0.38       | 46.8            | 0.39     | 0.9     |
| Functional disability          | 10.5        | 8.7        | 0.19       | 9.0             | 0.21     | 0.4     |
| Excellent/very good health     | 56.3        | 58.9       | 0.39       | 57.4            | 0.40     | 1.5     |
| Obese                          | 32.2        | 31.6       | 0.34       | 32.0            | 0.36     | 0.4     |
| Current smoker                 | 13.8        | 13.3       | 0.24       | 14.0            | 0.26     | 0.7     |
## Comparison of Select Child Estimates Using the Old and New Sample Child Weight: 2019

| Outcome                                      | Base Weight | Old Weight | New Weight | |New – Old| |
|----------------------------------------------|-------------|------------|------------|---|----------------|
| Uninsured                                    | 4.2         | 4.3        | 4.3        | 0.30 | 0.30 | 0.1 |
| ER visit, past year                          | 17.9        | 18.2       | 18.0       | 0.52 | 0.53 | 0.2 |
| Flu vaccination, past year                   | 51.9        | 52.1       | 51.7       | 0.67 | 0.68 | 0.4 |
| Asthma episode, past year                    | 3.3         | 3.3        | 3.3        | 0.23 | 0.24 | 0.0 |
| Excellent/very good health                   | 87.8        | 87.6       | 87.5       | 0.45 | 0.46 | 0.1 |
| Stressful life event, ever                   | 17.2        | 16.6       | 16.0       | 0.49 | 0.47 | 0.6 |
| Current ADD/ADHD                             | 8.1         | 7.3        | 7.1        | 0.33 | 0.32 | 0.2 |
Moved from geography-based NR adjustments to adjustments within response propensity classes based on response models using rich auxiliary data

- Old approach was static and could not adjust for changing patterns of nonresponse over time
- New approach will update response models annually, ensuring NR adjustments continue to perform well

Moved from weight calibration via post-stratification to raking

- Raking is more flexible and allows additional dimensions (e.g., education, MSA status, Census division)
Thank you!

Jim Dahlhamer
jdahlhamer@cdc.gov

- Coming to the NHIS website in late September: