shadac
STATE HEALTH ACCESS DATA ASSISTANCE CENTER

Robert Wood Johnson Foundation

AUGUST 2023

Four Methods for Calculating Income as a Percent of the Federal Poverty Guideline (FPG) in the Behavioral Risk Factor Surveillance System (BRFSS)

Author

Robert Hest, MPP

Research Fellow, State Health Access Data Assistance Center

Categorical vs. Continuous Income Measures

Categorical income measures ask respondents to report their income in several discrete categories, for example: "Is your annual household income from all sources…Less than $10,000, $10,000 to less than $15,000...$200,000 or more, etc."

Continuous income measures ask respondents to report their income as an exact dollar figure, for example: "What was this person's total income during the past 12 months? $__,___,___.00"

Stay updated by visiting us at **www.shadac.org** *or* **click here** *to join our mailing list.*

INTRODUCTION

Many surveys relevant to researchers evaluating the effects of health reform have income measures that lack needed specificity to precisely define the relevant eligible populations, such as those with incomes that make them eligible for Medicaid. Nevertheless, because these surveys are often the best (or only) source for a rich set of measures of health outcomes and health behaviors, researchers have employed various methods to best use the available income measures to examine potentially eligible populations.

The U.S. Centers for Disease Control and Prevention's (CDC) Behavioral Risk Factor Surveillance System (BRFSS) survey is one such resource; an annual household telephone survey of civilian non-institutionalized adults age 18 years or older that asks respondents about health behaviors, chronic health conditions, and the use of preventive health services. The BRFSS has a relatively large sample size, interviewing more than 400,000 respondents annually, and allows researchers to produce estimates for all 50 states and D.C.

The BRFSS' income measure asks respondents to report their total annual household income within 11 possible categories. Because these categories do not align with the federal poverty guideline (FPG) thresholds used to determine eligibility for programs such as Medicaid expansion (up to 138% FPG) or the Affordable Care Act's (ACA) cost sharing reductions (up to 250% FPG) or premium tax credits (up to 400% FPG), this creates a problem for researchers who want to use the BRFSS to study health reform.*

To deal with this issue, researchers have typically chosen to assign a continuous income to the respondent based on the categorical income measure, choosing either the lower bound of each category,[1] the upper bound of each category,[2] or the midpoint of each category.[3] There is no clear consensus in the literature about which approach to assigning continuous income from categorical values is best, and we propose that the most appropriate method depends in part on the analytic issue at hand.

In this brief, we first outline how each method impacts the income distribution in the BRFSS overall and by state. We then use the Current Population Survey (CPS) to evaluate the impact of using different methods to assign continuous income from a categorical income variable. We chose the CPS because the survey is used broadly to report on income and contains both a categorical and a continuous income variable. As a result, we can compare the impacts of different strategies of assigning continuous income from a categorical variable to actual, continuous income from the same data source. We then summarize our findings from the CPS and their implications for evaluating the impact of different health reform policies (e.g., Medicaid expansion) on health outcomes in the BRFSS.

* The BRFSS's household income measure is also problematic because eligibility for these programs is determined at the family level, not the household level, and the BRFSS does not have a measure of family income. Additionally, research shows that omnibus household income measures such as the one used in the BRFSS lead respondents to underreport their household income relative to aggregated household income measures. This leads to over-estimates of the poverty rate. See Davern et al., "The Effect of Income Question Design in Health Surveys on Family Income, Poverty and Eligibility Estimates."

Assigning Continuous Income in the BRFSS

Table 1 shows the BRFSS income categories and the relevant cut points for each method of assigning continuous income. As Table 1 demonstrates, if a respondent reported that their income was between $10,000 and $15,000, they would be assigned an income of $10,000 using the lower bound method, $15,000 using the upper bound method, or $12,500 using the midpoint method. The uniform distribution method randomly assign respondents an income value within their specified income category (e.g., $0-9,999) to create a uniform distribution of income within each category.

Income categories become wider as they increase. As a result, assignments in higher income categories are less precise. The top category, "$200,000 or more" has no upper bound, so we assigned an artificial upper bound of $250,000 to match the category size of the next lowest category (income $150,000 to less than $200,000).

Note that new categories have been added to the income question in the 2021 BRFSS with the category $75,000 or more being replaced by the categories $75,000 to less than $100,000; $100,000 to less than $150,000; $150,000 to less than $200,000; and $200,000 or more

Table 1. BRFSS Income Categories and Four Methods of Assigning Income

BRFSS Category	Category Size ($)	Lower Bound ($)	Midpoint ($)	Upper Bound ($)	Uniform Distribution ($)
Less than $10,000	10,000	0	5,000	10,000	0-9,999
$10,000 to less than $15,000	5,000	10,000	12,500	15,000	10,000-14,999
$15,000 to less than $20,000	5,000	15,000	17,500	20,000	15,000-19,999
$20,000 to less than $25,000	5,000	20,000	22,500	25,000	20,000-24,999
$25,000 to less than $35,000	10,000	25,000	30,000	35,000	25,000-34,999
$35,000 to less than $50,000	15,000	35,000	42,500	50,000	35,000-49,999
$50,000 to less than $75,000	25,000	50,000	62,500	75,000	50,000-74,999
$75,000 to less than $100,000	25,000	75,000	87,500	100,000	75,000-99,999
$100,000 to less than $150,000	50,000	100,000	125,000	150,000	100,000-149,999
$150,000 to less than $200,000	50,000	150,000	175,000	200,000	150,000-199,999
$200,000 or more	∞	200,000	225,000	250,000	200,000-249,999

Source: SHADAC analysis of the 2021 Behavioral Risk Factor Surveillance System.

Comparing Results of Four Income Assignment Methods in the BRFSS

Table 2 shows the results of potentially implementing each of these four income assignment methods in the BRFSS. As shown, there is substantial variation in the distribution of income across methods, particularly in how the methods affect estimates at the lower- and upper-most income categories. Not surprisingly, the lower bound method results in the largest share of the population with incomes below 100% FPG, while the upper bound method skews the income distribution toward the 401%+ FPG category. The uniform distribution and midpoint method are similar in both their impact on the overall income distribution and difference from the lower and upper bound methods.

Table 2. Distribution of Income (%FPG) by Income Assignment Method, 2021

Income Category (% FPG)	Lower Bound (%)	Uniform Distribution (%)	Midpoint (%)	Upper Bound (%)
0-100	17.8	13.6	13.3	10.9
101-138	8.5	8.4	10.8	6.4
139-250	28.3	22.6	23.6	19.3
251-400	27.9	30.2	31.9	29.4
401+	17.5	25.2	20.4	34.0
Total	100.0	100.0	100.0	100.0

Source: SHADAC analysis of the 2021 Behavioral Risk Factor Surveillance System; Universe: Civilian Non-institutionalized Adults. Florida did not collect sufficient data to be included in the 2021 BRFSS sample. National data do not include Florida.

Table 3 shows the percent of adults with incomes at or below 100% FPG by state, comparing each income imputation method, again using data from the 2021 BRFSS. The table shows substantial variation in state-level poverty rates and substantial state variation in the difference in poverty rates produced by each income imputation method.

The lower bound produces the highest rate of poverty across all states, followed in descending order by the uniform distribution, midpoint, and upper bound methods. In all but D.C., the uniform distribution method produces a significantly lower poverty rate compared to the lower bound method; the midpoint method produces a significantly lower poverty rate compared to the uniform distribution only at the national level; and in 46 states, the upper bound method produces a significantly lower poverty rate compared to the midpoint method.

There is also substantial variation in the size of the difference between the lower-bound and upper-bound estimates. D.C. has the smallest difference at 3.4 percentage points (pp), and New Mexico has the largest difference at 10.0 pp. The national average is 6.9 pp.

Table 3. Percent of Adults with Incomes at or Below 100% FPG by Income Assignment Method and State, 2021

State	Lower Bound (%)	Uniform Distribution (%)	Midpoint (%)	Upper Bound (%)	pp Difference Upper vs. Lower
Alabama	18.6	13.5*	13.3	10.5*	-8.1
Alaska	19.5	13.2*	14.5	11.8*	-7.7
Arizona	20.3	14.7*	15.1	12.2*	-8.1
Arkansas	22.3	17.3*	18.0	13.9*	-8.4
California	25	20.7*	21.4	18.5*	-6.5
Colorado	12.9	9.2*	9.5	7.1*	-5.8
Connecticut	13.6	9.6*	10.0	7.4*	-6.2
Delaware	15.5	11.1*	11.4	8.8*	-6.7
District of Columbia	12.5	10.1	10.2	9.1	-3.4
Georgia	20.4	14.5*	15.1	11.9*	-8.5
Hawaii	20.2	14.6*	13.9	12.1*	-8.1
Idaho	15.7	10.4*	10.7	7.7*	-8.0
Illinois	13.5	10.1*	10.5	8.1	-5.4
Indiana	17.2	12.8*	12.4	9.9*	-7.3
Iowa	12.8	9.4*	9.4	7.1*	-5.7
Kansas	13.8	9.3*	9.6	7.0*	-6.8
Kentucky	18.8	13.1*	13.8	10.5*	-8.3
Louisiana	23.2	17.5*	18.5	14.8*	-8.4
Maine	12.6	8.6*	8.9	5.9*	-6.7
Maryland	13.9	10.6*	10.7	9.0*	-4.9

Massachusetts	12	8.2*	8.7	6.5*	-5.5
Michigan	15.4	10.8*	11.2	8.5*	-6.9
Minnesota	10.6	7.5*	7.7	5.7*	-4.9
Mississippi	23.4	16.9*	17.9	13.6*	-9.8
Missouri	17.9	13.9*	14.3	11.6*	-6.3
Montana	14.6	10.3*	10.0	7.3*	-7.3
Nebraska	14.2	10.1*	10.1	7.7*	-6.5
Nevada	19.2	13.3*	13.4	10.0*	-9.2
New Hampshire	8.9	5.8*	6.2	3.9*	-5.0
New Jersey	14.3	10.5*	10.9	8.9*	-5.4
New Mexico	24.6	18.2*	18.5	14.6*	-10.0
New York	18.8	14.1*	14.6	11.8*	-7.0
North Carolina	17.8	12.2*	12.8	9.9*	-7.9
North Dakota	10.4	7.6*	7.3	5.3*	-5.1
Ohio	15.3	11.2*	11.1	8.5*	-6.8
Oklahoma	18.9	14.1*	14.1	10.9*	-8.0
Oregon	22.1	16.5*	17.2	13.1*	-9.0
Pennsylvania	12.2	9.1*	8.8	6.7*	-5.5
Rhode Island	15.1	11.1*	11.1	8.6*	-6.5
South Carolina	18.9	13.0*	13.9	10.3*	-8.6
South Dakota	13.4	9.0*	9.1	6.9	-6.5
Tennessee	18	13.2*	13.0	10.5*	-7.5
Texas	21.9	16.0*	16.3	13.1*	-8.8
Utah	12.6	8.7*	8.8	6.9*	-5.7
Vermont	11.5	7.2*	7.6	5.1*	-6.4
Virginia	11.9	8.4*	8.6	6.7*	-5.2
Washington	12.3	9.0*	8.9	7.0*	-5.3
West Virginia	18.4	13.1*	13.7	10.1*	-8.3
Wisconsin	10.6	7.8*	7.8	6.2	-4.4
Wyoming	13.5	8.9*	9.8	6.6*	-6.9
United States	17.8	13.3*	13.6*	10.9*	-6.9

** Difference from column to the left statistically significant at 95% level.*
Source: SHADAC analysis of the 2021 Behavioral Risk Factor Surveillance System; Universe: Civilian Non-institutionalized Adults.
Florida did not collect sufficient data to be included in the 2021 BRFSS sample, therefore it is not included in this table. National data do not include Florida.

Results in the CPS of Four Methods of Assigning Income

Although it is helpful to understand the variation across these methods within the BRFSS, we also wanted to compare the outcomes of each method to the outcome produced by a measure of continuous income. The CPS provides a helpful data source to make this comparison because it contains both categorical and continuous income information. To evaluate the impact of each method, we first recoded the CPS's categorical income category to match the BRFSS categories as closely as possible. (CPS's categorical income variable top category is income at or above $150,000, which is less than BRFSS' top category of income at or above $200,000.) We then assigned a continuous income value based on each of the lower, midpoint, upper bound, and uniform distribution methods as described previously. Finally, we assigned FPG based on family size, calculated income as a percent of FPG, and compared the results to FPG calculated based on the actual continuous income measure observed in the CPS data set, which for the purposes of this analysis will be considered the "true" or reference income distribution.

Table 4 shows the results of our analysis by FPG categories typically used to determine eligibility for Medicaid or ACA coverage subsidies. As the table demonstrates, similar to our findings in the BRFSS, there is substantial variation across methods on the

overall income distribution. However, there is also variation in how each method measures up against the "true" distribution based on continuous income.

Overall accuracy: In terms of coming closest to matching the reference distribution across the majority of FPG categories, the uniform distribution method performs best. The uniform distribution method matches or outperforms the other methods in estimating the share of the population in all but the 101-138% FPG category.

Accuracy for Medicaid & Cost Sharing-eligible Categories: The uniform distribution performs best in estimating the 0-100% FPG population, though the upper-bound method performs better in estimating the 101-138% FPG category and performs similarly to the uniform distribution in estimating the 139-250% FPG category.

Table 4. Distribution of Income (% FPG) by Income Assignment Method

Income Assignment Method		Percent	Difference from Reference (Percentage Point)
Lower Bound	**0-100% FPG (Medicaid)**	14.6	3.7
	101-138% FPG (Medicaid)	7.4	1.6
	139-250% FPG (Cost Sharing)	26.6	7.9
	251-400% FPG (Premium Subsidies)	29.9	9.7
	401%+ FPG	21.4	-22.8
	Total	100.0	
Uniform Distribution	**0-100% FPG**	11.4	0.5
	101-138% FPG	6.7	0.8
	139-250% FPG	20.5	1.7
	251-400% FPG	30.8	10.6
	401%+ FPG	30.7	-13.6
	Total	100.0	
Midpoint	**0-100% FPG**	10.1	-0.9
	101-138% FPG	6.2	0.4
	139-250% FPG	20.8	2.0
	251-400% FPG	33.6	13.4
	401%+ FPG	29.3	-15.0
	Total	100.0	
Upper Bound	**0-100% FPG**	9.3	-1.7
	101-138% FPG	5.3	-0.5
	139-250% FPG	17.2	-1.6
	251-400% FPG	20.8	0.6
	401%+ FPG	47.5	3.2
	Total	100.0	
Reference Income	0-100% FPG	10.9	
	101-138% FPG	5.8	
	139-250% FPG	18.8	
	251-400% FPG	20.2	
	401%+ FPG	44.3	
	Total	100.0	

Definition: 2021 family income as a percent of the 2021 federal poverty guidelines (FPG) for the civilian, non-institutionalized population of adults age 18 or older.
Source: SHADAC analysis of the 2022 Current Population Survey's Annual Social and Economic Supplements (CPS-ASEC) public use microdata file..

CONCLUSION

As these findings demonstrate, if the goal is to assign income to most closely match the overall distribution of income derived from actual continuous income, the best overall method for assigning continuous income from categorical income is the uniform distribution method. Except in specific cases, the other evaluated methods generally failed to outperform the uniform distribution method in overall accuracy.

This finding differs from our previous analysis of the BRFSS' former income categories, which found that the accuracy of the methods analyzed here varied substantially by %FPG income category, with the upper-bound method performing best overall but failing to accurately reflect the share of the population with Medicaid-eligible incomes. This appears to no longer be the case with BRFSS' extended income categories, which now more precisely reflect the income of higher-income respondents.

References

Benitez, Joseph A., and Eric E. Seiber. (2018). US Health Care Reform and Rural America: Results From the ACA's Medicaid Expansions. *Journal of Rural Health*, *34*(2), 213–22. doi:10.1111/jrh.12284.

Centers for Disease Control and Prevention (CDC). (2015). BRFSS Statistical Brief on the Health Care Access Module, 2013 and 2014 [PDF file]. Retrieved from https://www.cdc.gov/brfss/data_documentation/pdf/2013-2014_hcs.pdf.

Davern, Michael, Holly Rodin, Timothy J. Beebe, and Kathleen Thiede Call. (2005). The Effect of Income Question Design in Health Surveys on Family Income, Poverty and Eligibility Estimates. *Health Services Research*, *40*(5), 1534-1552. doi:10.1111/j.1475-6773.2005.00416.x.

Hawaii Health Data Warehouse. (n.d.) HHDW BRFSS Poverty Level Methodology [PDF file] Retrieved from http://www.hhdw.org/wp-content/uploads/BRFSS-Poverty-Level-Methodology.pdf.

Levy, Helen, and David Meltzer. (2007). The Impact of Health Insurance on Health. *Annual Review of Public Health*, *29*(1), 399-409. doi:10.1146/annurev.publhealth.28.021406.144042.

Lopez, Russ. (2004). Income Inequality and Self-Rated Health in US Metropolitan Areas: A Multi-Level Analysis. *Social Science and Medicine*, *59*(12), 2409-2419. doi:10.1016/j.socscimed.2004.03.033.

Nasseh, Kamyar, and Marko Vujicic. (2014). The Effect of Growing Income Disparities on U.S. Adults' Dental Care Utilization. *Journal of the American Dental Association*, *145*(5), 435-442. doi:10.14219/jada.2014.1.

Sabik, Lindsay M., Wafa W. Tarazi, and Cathy J. Bradley. (2015). State Medicaid Expansion Decisions and Disparities in Women's Cancer Screening. *American Journal of Preventive Medicine*, *48*(1), 98-103. doi:10.1016/j.amepre.2014.08.015.

Simon, Kosali, Aparna Soni, and John Cawley. (2017). The Impact of Health Insurance on Preventive Care and Health Behaviors: Evidence from the First Two Years of the ACA Medicaid Expansions. *Journal of Policy Analysis and Management*, *36*(2), 390-417. doi:10.1002/pam.21972.

State Health Access Data Assistance Center. (2019). FPG vs. FPL: What's the Difference? *SHADAC Blog*, Retrieved from https://www.shadac.org/news/fpg-vs-fpl-whats-difference.

Yue, Dahai, Petra W. Rasmussen, and Ninez A. Ponce. (2018). Racial/Ethnic Differential Effects of Medicaid Expansion on Health Care Access. *Health Services Research*, *53*(5), 3640-3656. doi:10.1111/1475-6773.12834.

Appendix: Implementing Income Imputation Methods in the BRFSS

Each of these strategies were implemented using microdata from the 2021 BRFSS public use file, using data from respondents residing in the contiguous United States, Alaska, and Hawaii. All analysis was conducted using Stata v18.1.

Assigning Continuous Income

The first step in the analysis is to assign each respondent a continuous income value based on their response to the BRFSS categorical income measure. Respondents with missing income information were dropped from the analysis. The code below shows how continuous incomes were assigned using the midpoint method. The same code was used for the lower-bound and upper-bound strategies, but using the income values shown in Table 1.

```
generate income_mid = .
replace income_mid = 5000 if income3 == 1
replace income_mid = 12500 if income3 == 2
replace income_mid = 17500 if income3 == 3
replace income_mid = 22500 if income3 == 4
replace income_mid = 30000 if income3 == 5
replace income_mid = 42500 if income3 == 6
replace income_mid = 62500 if income3 == 7
replace income_mid = 87500 if income3 == 8
replace income_mid = 112500 if income3 == 9
replace income_mid = 175000 if income3 == 10
replace income_mid = 225000 if income3 == 11
```

The code below shows how the uniform distribution strategy was implemented. The STATA random number function "runiform" was used to randomly assign respondents an income value from the uniform distribution based on their relevant income category. Using the "set seed" function allows the same results to be later reproduced. In this case the seed was set to a random, four-digit number generated by www.random.org.

```
set seed 3849
gen income_unif = .
replace income_unif = runiform(0,9999.99) if income3 == 1
replace income_unif = runiform(10000,14999.99) if income3 == 2
replace income_unif = runiform(15000,19999.99) if income3 == 3
replace income_unif = runiform(20000,24999.99) if income3 == 4
replace income_unif = runiform(25000,34999.99) if income3 == 5
replace income_unif = runiform(35000,49999.99) if income3 == 6
replace income_unif = runiform(50000,74999.99) if income3 == 7
replace income_unif = runiform(75000,99999.99) if income3 == 8
replace income_unif = runiform(100000,149999.99) if income3 == 9
replace income_unif = runiform(150000,199999.99) if income3 == 10
replace income_unif = runiform(200000,250000) if income3 == 11
```

Calculating Household Size

An individual's FPG is based on their family size. Because the BRFSS lacks a measure of family size, this analysis uses household size to determine the FPG.

The BRFSS has three variables that are combined to determine household size: HHADULT (the number of adults in the household for cellphone respondents), NUMADULT (the number of adults in the household for landline respondents), and CHILDREN (the number of children under 18 years of age in the household for all respondents). Observations with missing information for any of these variables or with more than 14 individuals in the household were dropped from the analysis.

The code below first cleans the HHADULT and CHILDREN variables, setting non-responses to "missing" and recoding the "no children" response to "0" for the CHILDREN variable. The code then sums each respondent's NUMADULT, cleaned HHADULT, and cleaned CHILDREN variables, setting the resulting variable to "missing" if any of the three input variables are missing. This variable is then set to "missing" if it is equal to 0 or if is greater than 14 or if either NUMADULT or HHADULT_CLEAN are missing.

```
ggen hhadult_clean = .
replace hhadult_clean = hhadult if hhadult != 77 & hhadult != 99

gen children_clean = .
replace children_clean = children if children != 99
replace children_clean = 0 if children == 88

egen hh_size = rowtotal(numadult hhadult_clean children_clean), missing
replace hh_size = . if hh_size == 0 | hh_size > 14
replace hh_size = . if hhadult_clean == . & numadult == .
```

Calculating Income as a Percent of the Federal Poverty Guideline (FPG)

In addition to family size, an individual's FPG is also determined by their state of residence. The FPG is higher in Alaska and Hawaii than it is in the 48 contiguous states and D.C.

The code below creates the STATEGROUP and YEAR variable to merge a base and increment FPG variable for each respondent in the dataset. These FPG base and increment variables are contained in a separate dataset based on the Federal Poverty Guidelines (FPG) released by the U.S. Department of Health and Human Services.

```
gen stategroup = .
replace stategroup = 1 if _STATE != 2 & _STATE != 15 & !mi(_STATE)
replace stategroup = 2 if _STATE == 2 // Alaska
replace stategroup = 3 if _STATE == 15 // Hawaii

gen year = 2021

merge m:1 year stategroup using "C:\...\HHS Poverty Guidelines.dta"
drop if _merge == 2
drop _merge year stategroup
```

The section of code below uses the merged base and increment variables to create a FPG for each respondent based on their state and household size. The FPG is created by adding the base FPG and an additional increment FPG for each additional household member.

```
gen fpg_guideline = .
replace fpg_guideline = base + (increment * (hh_size - 1)) if !mi(hh_size)
```

The code below produces income as a percent of FPG for each of the four methods of imputing continuous income in the BRFSS by dividing income by the FPG and multiplying the result by 100.

```
gen fpg_low = 100 * (income_low / fpg_guideline)
gen fpg_mid = 100 * (income_mid / fpg_guideline)
gen fpg_up = 100 * (income_up / fpg_guideline)
gen fpg_unif = 100 * (income_unif / fpg_guideline)
```

These FPG variables can then be used to create policy-relevant categorical variables; for example, for individuals at or below 100% FPG.

[1] Nasseh and Vujicic, "The Effect of Growing Income Disparities on U.S. Adults' Dental Care Utilization."

[2] Levy and Meltzer, "The Impact of Health Insurance on Health"; Simon, Soni, and Cawley, "The Impact of Health Insurance on Preventive Care and Health Behaviors: Evidence from the First Two Years of the ACA Medicaid Expansions"; Yue, Rasmussen, and Ponce, "Racial/Ethnic Differential Effects of Medicaid Expansion on Health Care Access."

[3] Benitez and Seiber, "US Health Care Reform and Rural America: Results From the ACA's Medicaid Expansions"; Centers for Disease Control and Prevention (CDC), "BRFSS Statistical Brief on the Health Care Access Module, 2013 and 2014"; Hawaii Health Data Warehouse, "HHDW BRFSS Poverty Level Methodology"; Lopez, "Income Inequality and Self-Rated Health in US Metropolitan Areas: A Multi-Level Analysis"; Sabik, Tarazi, and Bradley, "State Medicaid Expansion Decisions and Disparities in Women's Cancer Screening."