

Steps in Quantitative Research

Example research questions

Research question 1: what are the racial/ethnic disparities in *health-related quality of life (HRQOL)* among the general population and among people with low SES (education and/or income) particularly?

Research question 2: What are the racial/ethnic disparities in *health care access* among the general population and among people with low SES (education and/or income) particularly?

1. Obtain data

- Can the data be used to address the research questions and hypotheses?
- How were the data collected?
 - If random sampling, what was the sampling technique?
 - If intervention study, what was the study design?
- What was the sampling unit of the data? What was the sample size?
- Do we have enough statistical power to test the research hypotheses if required?

BRFSS 2021 data

- Can BRFSS 2021 data be used to address the research questions?
- For health-related quality of life, BRFSS 2021 has
 - the days of physical health not good (PHYSHLTH)
 - the days of mental health not good (MENTHLTH)
 - the days poor physical or mental health keep you from doing your usual activities (POORHLTH).
- For health care access, BRFSS 2021 has
 - visited a doctor within the past year (CHECKUP1)
 - could not see a doctor when needed in the past year (MEDCOST1).
- Independent variables: racial/ethnic groups, education, income, health insurance coverage, and other basic demographic variables.

BRFSS 2021 data

- BRFSS 2021 is a national complex samples telephone survey of noninstitutionalized U.S. residents aged 18 years and older (the population).
- The total sample size is 438,693 in 2,140 strata with a median response rate of 44% (ranged from 23.5% to 60.5%).
- Florida was unable to collect enough BRFSS data in 2021 to meet the minimum requirements for inclusion in the 2021 BRFSS public-use data set (undercoverage).

Research hypotheses

Research question 1: what are the racial/ethnic disparities in health-related quality of life (HRQOL) among the general population and among people with low SES (education and/or income) particularly?

- Hypothesis 1.1: *The average # of days of physical health not good* differs among racial/ethnic groups.
- Hypothesis 1.2: This racial/ethnic disparity deepens in people with low SES.

Research question 2: What are the racial/ethnic disparities in health care access among the general population and among people with low SES (education and/or low income) particularly?

- Hypothesis 2.1: *The odds of visited a doctor within the past year* differs among racial/ethnic groups.
- Hypothesis 2.2: This racial/ethnic disparity deepens in people with low SES.

Power and sample size analysis

- Formal power and sample size analysis is not necessary for exploratory data analysis. Describe the data generation process such as inclusion/exclusion criteria.
- For each research hypothesis, hypothesized values of parameters are needed. For example:
 - Hypothesis 1.1: *The average # of days of physical health not good* differs among racial/ethnic groups.
 - For quantitative data, we need mean and standard deviation of PHYSHLTH for each racial/ethnic group.
 - Hypothesis 2.1: *The odds of visited a doctor within the past year* differs among racial/ethnic groups.
 - For categorical data, we need % people visited a doctor within the past year for each racial/ethnic group.
- For subgroup analyses, the sample size requirement is higher.

2. Data cleaning

- Know how the data were originally collected?
- Go through the data and obtain summary statistics to verify the data. Do some data cleaning if necessary.
- Recode the data if necessary to fit the research questions and hypotheses.
- Pay attention to missing data. How do we handle missing data? Is imputation needed.

BRFSS 2021 data

Label: Number of Days Physical Health Not Good

Section Name: Healthy Days

Core Section Number: 2

Question Number: 1

Column: 102-103

Type of Variable: Num

SAS Variable Name: PHYSHLTH

Question Prologue:

Question: Now thinking about your physical health, which includes physical illness and injury, for how many days during the past 30 days was your physical health not good?

Use it as quantitative data or categorize it?

Value	Value Label	Frequency	Percentage	Weighted Percentage
1 - 30	Number of days	141,403	32.23	31.47
88	None → 0	287,796	65.60	66.53
77	Don't know/Not sure → Missing	7,898	1.80	1.66
99	Refused → data	1,593	0.36	0.34
BLANK	Not asked or Missing →	3	.	.

BRFSS 2021 data

Label: Length of time since last routine checkup

Section Name: Health Care Access

Core Section Number: 3

Question Number: 4

Column: 112

Type of Variable: Num

SAS Variable Name: CHECKUP1

Question Prologue:

Question: About how long has it been since you last visited a doctor for a routine checkup? [A routine checkup is a general physical exam, not an exam for a specific injury, illness, or condition.]

Use it as ordinal data or dichotomize it?

Value	Value Label	Frequency	Percentage	Weighted Percentage
1	Within past year (anytime < 12 months ago)	337,289	76.89	72.51
2	Within past 2 years (1 year but < 2 years ago)	51,271	11.69	13.78
3	Within past 5 years (2 years but < 5 years ago)	23,427	5.34	6.68
4	5 or more years ago	18,490	4.21	4.83
7	Don't know/Not sure → Missing data	5,233	1.19	1.32
8	Never → Combine with 4	2,311	0.53	0.72
9	Refused → Missing data	670	0.15	0.17
BLANK	Not asked or Missing → Missing data	2	.	.

BRFSS 2021 data

Label: Imputed race/ethnicity value
Section Name: Weighting Variables
Module Number: 1
Question Number: 12
Column: 1471-1472
Type of Variable: Num
SAS Variable Name: _IMPRACE

Question Prologue:

Question: Imputed race/ethnicity value (This value is the reported race/ethnicity or an imputed race/ethnicity, if the respondent refused to give a race/ethnicity. The value of the imputed race/ethnicity will be the most common race/ethnicity response for that region of the state)

For a sample size consideration, are some minority groups too small for data analysis?

Value	Value Label	Frequency	Percentage	Weighted Percentage
1	White, Non-Hispanic	332,222	75.73	62.08
2	Black, Non-Hispanic	33,132	7.55	11.64
3	Asian, Non-Hispanic	11,557	2.63	5.88
4	American Indian/Alaskan Native, Non-Hispanic	7,410	1.69	0.96
5	Hispanic	38,688	8.82	17.19
6	Other race, Non-Hispanic	15,684	3.58	2.26

Missing data

- With about 2% missing data, it should not be a concern here. Analyzing complete cases (cases with no missing data) should not prohibit us from generalize our results to the population.
- If less than 10% data are missing on each variable, we can impute the missing data by methods such as hot deck, mean imputation, or regression.
- Often, multiple imputations are preferred to account for the extra variations from the missing data.



• But few good methods can help effectively deal with excessive missing data.

BRFSS 2021 data

Label: Computed income categories
Section Name: Calculated Variables
Module Number: 9
Question Number: 24
Column: 2008
Type of Variable: Num
SAS Variable Name: _INCOMG1
Question Prologue:
Question: Income categories

Value	Value Label	Frequency	Percentage	Weighted Percentage
1	Less than \$15,000 Notes: INCOME3=1,2	22,408	5.11	6.02
2	\$15,000 to < \$25,000 Notes: INCOME3=3,4	36,031	8.21	8.32
3	\$25,000 to < \$35,000 Notes: INCOME3=5	43,893	10.01	9.75
4	\$35,000 to < \$50,000 Notes: INCOME3=6	48,339	11.02	10.12
5	\$50,000 to < \$100,000 Notes: INCOME3=7,8	107,246	24.45	22.58
6	\$100,000 to < \$200,000 Notes: INCOME3=9,10	67,411	15.37	15.68
7	\$200,000 or more Notes: INCOME3=11	18,952	4.32	5.28
9	Don't know/Not sure/Missing Notes: INCOME3=77, 99, or missing	94,413	21.52	22.25

Missing data!

Imputation

- Hot-deck/cold-deck: impute data from similar record(s) either from the same dataset or from a different dataset. May involve sorting the dataset by some variable(s) based on the observed data.
- Mean imputation: impute data with the mean of the variable based on all observed cases. This underestimates the variance and attenuates any correlations between the imputed variable and other variables.
- Regression imputation: impute data with predicted values from a regression model. This still underestimates the variance but may overestimate correlations with other variables.
- Multiple imputation: impute data with multiple randomly generated values (from a model). Multiple copies of imputed data will be generated and analyzed, and results will be pooled together.

Missing data mechanisms

- Missing completely at random: the reason the data are missing is completely random.
 - Example: simply forgot to provide responses to some survey questions.
- Missing at random: the cause of the missingness depends on some observed variable(s) for which data have been collected.
 - Example: failure to provide responses to some survey questions may be related to one's age and health conditions which are available.
- Missing not at Random: the cause of the missingness may depend on the value of the missing data itself so cannot be controlled for.
 - Example: not willing to provide income because it is very high (for a tax concern). A survey of an acute illness missed some respondents because they were ill at home. Or the variable(s) related to the missingness is simply not measured.

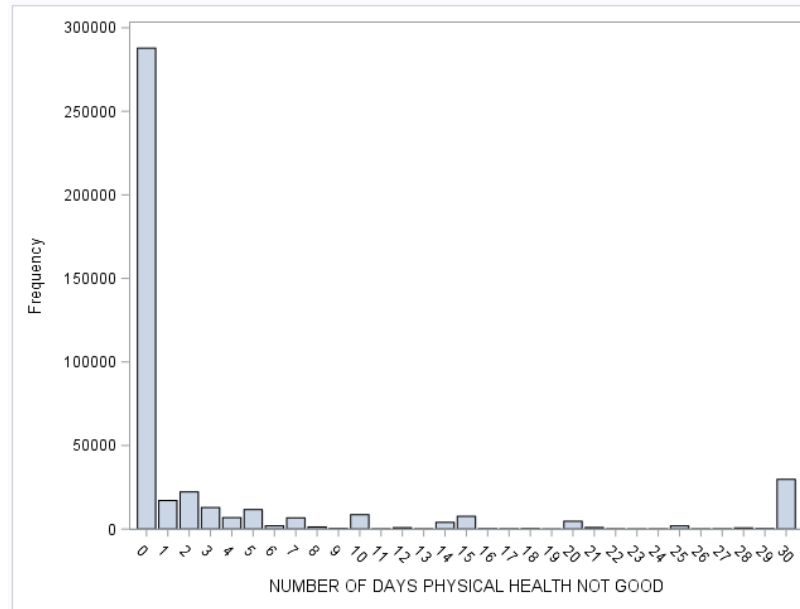
3. Data summary

- Obtain data summary plots and statistics for all data involved.
- Pay attention to the distributions of the data.
 - If any outliers, are they to be removed?
 - If strong skewness in the data, are remediation measures necessary?
 - Is sensitivity analysis necessary?
- Data summary will also help you choose appropriate statistical methods to analyze the data.

Assumptions in analysis

- Two-sample t-test
 - The two groups of data are independent and not too far from a normal distribution.
 - A larger sample size is required to deal with more skewed data.
- Analysis of variance
 - Each group of data are independent and not too far from a normal distribution.
 - Group variances are often assumed the same.
- Linear regression
 - The relationship between the response and predictor(s) are linear.
 - The errors are independent and not too far from a normal distribution.
 - The error variance is constant.
- Logistic regression
 - The data are independent, and the logit (log odds) is linearly related to the predictor(s).

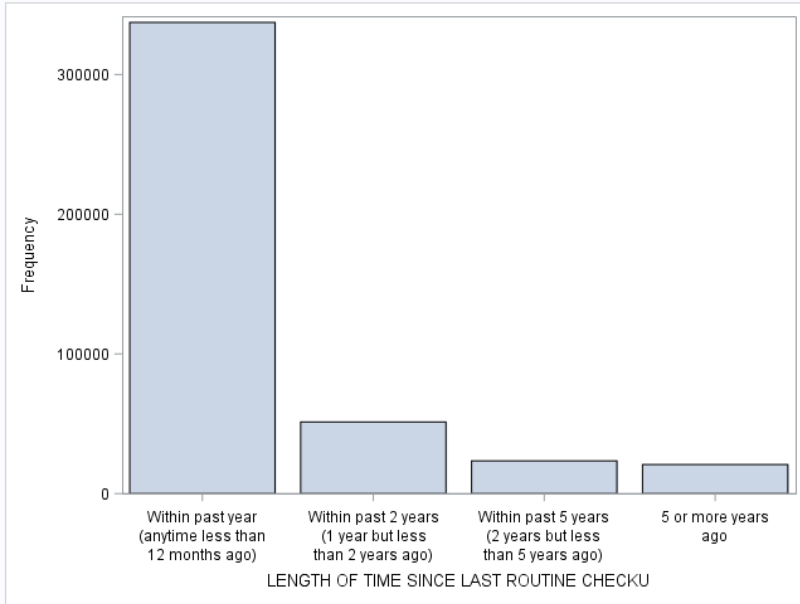
BRFSS 2021 data



Analysis Variable : PHYSHLTH NUMBER OF DAYS PHYSICAL HEALTH NOT GOOD

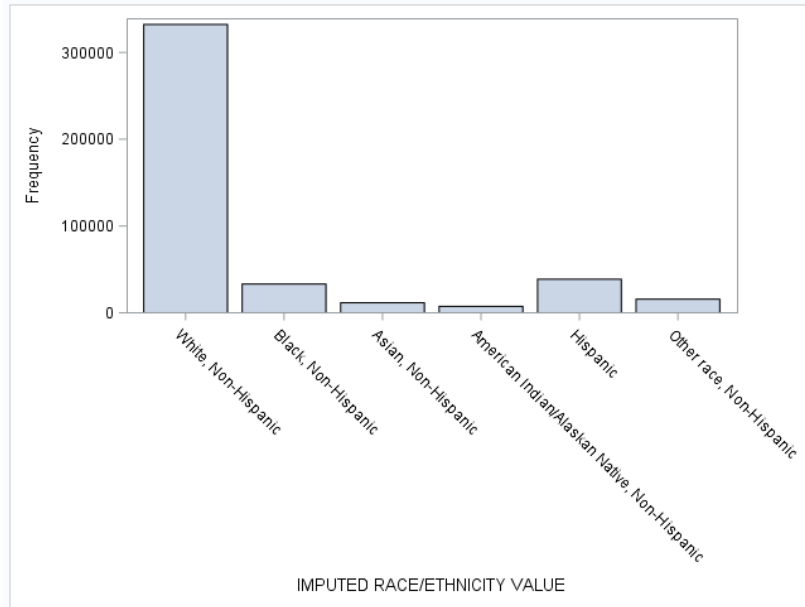
N	Mean	Std Dev	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
429199	3.80	8.35	0.00	0.00	0.00	2.00	30.00

BRFSS 2021 data



LENGTH OF TIME SINCE LAST ROUTINE CHECKUP				
CHECKUP1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Within past year (anytime less than 12 months ago)	337289	77.93	337289	77.93
Within past 2 years (1 year but less than 2 years ago)	51271	11.85	388560	89.78
Within past 5 years (2 years but less than 5 years ago)	23427	5.41	411987	95.19
5 or more years ago	20801	4.81	432788	100.00
Frequency Missing = 5905				

BRFSS 2021 data



IMPUTED RACE/ETHNICITY VALUE				
_IMPRACE	Frequency	Percent	Cumulative Frequency	Cumulative Percent
White, Non-Hispanic	332222	75.73	332222	75.73
Black, Non-Hispanic	33132	7.55	365354	83.28
Asian, Non-Hispanic	11557	2.63	376911	85.92
American Indian/Alaskan Native, Non-Hispanic	7410	1.69	384321	87.61
Hispanic	38688	8.82	423009	96.42
Other race, Non-Hispanic	15684	3.58	438693	100.00

4. Data analysis

- Choose appropriate statistical methods to address the research questions and hypotheses using available data.
- Conduct the analyses and investigate the results.
- All statistical models have assumptions. Verify those assumptions using diagnostic plots and statistics. If assumptions were violated, take necessary remediation measures.
- Is the chosen statistical model appropriate? Are the chosen variables necessary and enough to address the research questions and hypotheses. If not, revise the model and refine the variable selection.
- Run the revised analyses, investigate the results, and check model diagnostics again. Is it satisfactory? If not, revise the model again.

PHYSHLTH vs. race/ethnicity

- The days of physical health not good (PHYSHLTH) is strongly skewed with excessive number of zeros.
- It is of practical sense to compare the mean # of days physical health not good (as an exploratory analysis) among the racial/ethnic populations.
- But testing the hypothesis of any difference among the racial/ethnic populations using ANOVA might be problematic. The complex samples procedures (especially with bootstrap variance estimation) may help.

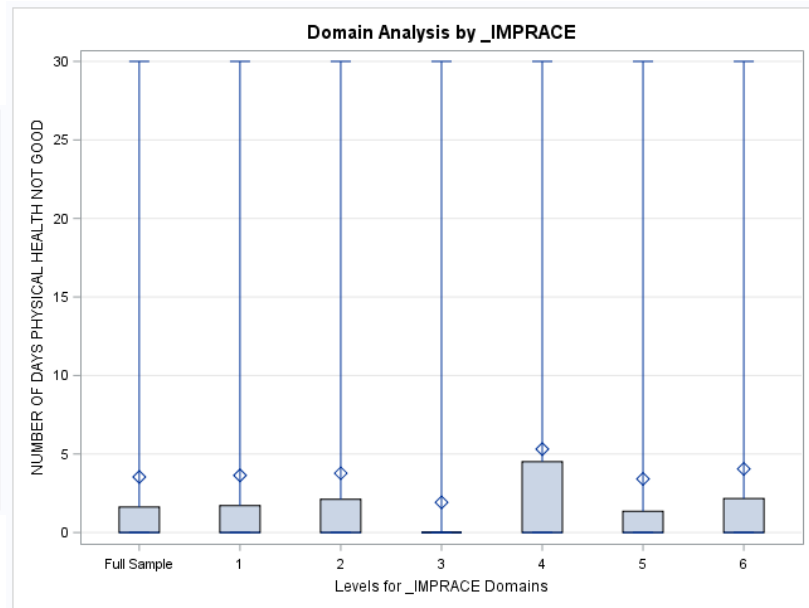


A crude analysis would be to dichotomize the data to zero or any days.

PHYSHLTH vs. race/ethnicity

Proc surveymeans

Statistics for _IMPRACE Domains							
IMPUTED RACE/ETHNICITY VALUE	Variable	Label	N	Mean	Std Error of Mean	95% CL for Mean	
White, Non-Hispanic	PHYSHLTH	NUMBER OF DAYS PHYSICAL HEALTH NOT GOOD	325385	3.637754	0.028419	3.58205387	3.69345315
Black, Non-Hispanic	PHYSHLTH	NUMBER OF DAYS PHYSICAL HEALTH NOT GOOD	32217	3.771516	0.081631	3.61152202	3.93151069
Asian, Non-Hispanic	PHYSHLTH	NUMBER OF DAYS PHYSICAL HEALTH NOT GOOD	11351	1.919175	0.147421	1.63023353	2.20811552
American Indian/Alaskan Native, Non-Hispanic	PHYSHLTH	NUMBER OF DAYS PHYSICAL HEALTH NOT GOOD	7188	5.309918	0.250908	4.81814602	5.80168946
Hispanic	PHYSHLTH	NUMBER OF DAYS PHYSICAL HEALTH NOT GOOD	37785	3.411069	0.075875	3.26235502	3.55978215
Other race, Non-Hispanic	PHYSHLTH	NUMBER OF DAYS PHYSICAL HEALTH NOT GOOD	15273	4.047943	0.154126	3.74586085	4.35002486



* SAS complained no sufficient memory for domain comparison!

PHYSHLTH vs. race/ethnicity

Proc surveyreg

Data Summary	
Number of Observations	429199
Sum of Weights	241109851
Weighted Mean of PHYSHLTH	3.53807
Weighted Sum of PHYSHLTH	853064005

Design Summary	
Number of Strata	2138
Number of Strata Collapsed	66
Number of Clusters	429199

Fit Statistics	
R-Square	0.003215
Root MSE	7.9831
Denominator DF	427126

Tests of Model Effects			
Effect	Num DF	F Value	Pr > F
Model	5	52.81	<.0001
Intercept	1	6110.30	<.0001
_IMPRACE	5	52.81	<.0001

Note: The denominator degrees of freedom for the F tests is 427126.

Estimated Regression Coefficients				
Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	3.6377535	0.01938710	187.64	<.0001
_IMPRACE American Indian/Alaskan Native, Non-Hispanic	1.6721642	0.19882037	8.41	<.0001
_IMPRACE Asian, Non-Hispanic	-1.7185790	0.13690935	-12.55	<.0001
_IMPRACE Black, Non-Hispanic	0.1337628	0.05927404	2.26	0.0240
_IMPRACE Hispanic	-0.2266849	0.06031160	-3.76	0.0002
_IMPRACE Other race, Non-Hispanic	0.4101893	0.12607538	3.25	0.0011
_IMPRACE White, Non-Hispanic	0.0000000	0.00000000	.	.

PHYSHLTH vs. race/ethnicity by income

Proc surveyreg with income level as a domain

Race/ethnicity	Income				
	All	<\$15K	\$15K-\$50K	\$50K-\$100K	>\$100K
American Indian/Alaskan Native, Non-Hispanic	1.67	-0.55	0.59	0.17	0.95
Asian, Non-Hispanic	-1.72	-6.05	-2.86	-0.90	-0.62
Black, Non-Hispanic	0.13	-2.07	-1.10	-0.51	0.19
Hispanic	-0.23	-3.77	-1.92	-0.28	0.07
Other race, Non-Hispanic	0.41	-0.11	-0.26	0.48	-0.20
White, Non-Hispanic	0.00	0.00	0.00	0.00	0.00

* SAS complained no sufficient memory for testing the interaction between race/ethnicity and income!

PHYSHLTH vs. race/ethnicity by income

Proc surveyreg with income level as a domain

Race/ethnicity	Income				
	All	<\$15K	\$15K-\$50K	\$50K-\$100K	>\$100K
American Indian/Alaskan Native, Non-Hispanic	1.98	-0.73	0.90	0.20	0.94
Asian, Non-Hispanic	-1.18	-4.89	-2.24	-0.58	-0.53
Black, Non-Hispanic	0.34	-2.22	-0.74	-0.39	0.18
Hispanic	0.29	-2.96	-1.13	0.08	0.19
Other race, Non-Hispanic	0.81	-0.03	0.23	0.75	-0.16
White, Non-Hispanic	0.00	0.00	0.00	0.00	0.00

* Analyses adjusted by age, sex, and health insurance coverage.

Dichotomized PHYSHLTH

Proc surveyfreq

Data Summary	
Number of Strata	2140
Number of Clusters	438693
Number of Observations	438693
Sum of Weights	246041640

Table of PHYSHLTH_2g							
PHYSHLTH_2g	Frequency	Weighted Frequency	Std Err of Wgt Freq	Percent	Std Err of Percent	95% Confidence Limits for Percent	
0	287796	163683774	520969	67.8876	0.1578	67.5784	68.1969
1	141403	77426077	406052	32.1124	0.1578	31.8031	32.4216
Total	429199	241109851	499400	100.0000			
Frequency Missing = 9494							

PHYSHLTH > 0 vs. race/ethnicity

Proc surveyfreq

RACE/ETHNICITY	PHYSHLTH>0	Weighted	Row percent	Std Err
White, Non-Hispanic	107848	49974343	33.34	0.17
Black, Non-Hispanic	11089	9253632	33.15	0.48
Asian, Non-Hispanic	2553	3380316	23.76	0.98
American Indian/Alaskan Native, Non-Hispanic	2756	891353	38.97	1.20
Hispanic	11750	12059279	29.17	0.47
Other race, Non-Hispanic	5407	1867154	34.32	0.83

Rao-Scott Chi-Square Test	
Pearson Chi-Square	1421.3381
Design Correction	6.0896
Rao-Scott Chi-Square	233.4024
DF	5
Pr > ChiSq	<.0001
F Value	46.6805
Num DF	5
Den DF	2135305
Pr > F	<.0001
Sample Size = 429199	

PHYSHLTH > 0 vs. race/ethnicity

Proc surveylogistic

Type 3 Analysis of Effects				
Effect	F Value	Num DF	Den DF	Pr > F
_IMPRACE	33.30	5	427057	<.0001

Odds Ratio Estimates			
Effect	Point Estimate	95% Confidence Limits	
_IMPRACE American Indian/Alaskan Native, Non-Hispanic vs White, Non-Hispanic	1.277	1.155	1.411
_IMPRACE Asian, Non-Hispanic vs White, Non-Hispanic	0.623	0.560	0.693
_IMPRACE Black, Non-Hispanic vs White, Non-Hispanic	0.992	0.948	1.037
_IMPRACE Hispanic vs White, Non-Hispanic	0.823	0.786	0.863
_IMPRACE Other race, Non-Hispanic vs White, Non-Hispanic	1.045	0.971	1.125

NOTE: The degrees of freedom in computing the confidence limits is 427061.

PHYSHLTH > 0 vs. race/ethnicity

Proc surveylogistic

Odds ratio	All	<\$15K	\$15K- \$50K	\$50K- \$100K	>\$100K
American Indian/Alaskan Native, Non-Hispanic vs White, Non-Hispanic	1.28	0.97	1.04	0.95	1.09
Asian, Non-Hispanic vs White, Non-Hispanic	0.63	0.34	0.49	0.78	0.80
Black, Non-Hispanic vs White, Non-Hispanic	0.99	0.84	0.79	0.84	0.93
Hispanic vs White, Non-Hispanic	0.82	0.46	0.56	0.89	1.01
Other race, Non-Hispanic vs White, Non-Hispanic	1.04	0.78	0.95	1.01	0.93

Joint Tests				
Effect	F Value	Num DF	Den DF	Pr > F
_IMPRACE	37.44	5	336318	<.0001
_income_4g	80.26	3	336320	<.0001
_IMPRACE*_income_4g	8.15	15	336308	<.0001

PHYSHLTH > 0 vs. race/ethnicity

Proc surveylogistic

Odds ratio	All	<\$15K	\$15K- \$50K	\$50K- \$100K	>\$100K
American Indian/Alaskan Native, Non-Hispanic vs White, Non-Hispanic	1.34	0.96	1.07	0.93	1.09
Asian, Non-Hispanic vs White, Non-Hispanic	0.66	0.36	0.52	0.78	0.82
Black, Non-Hispanic vs White, Non-Hispanic	1.00	0.82	0.80	0.84	0.92
Hispanic vs White, Non-Hispanic	0.87	0.47	0.59	0.93	1.04
Other race, Non-Hispanic vs White, Non-Hispanic	1.08	0.77	0.95	1.02	0.93

Type 3 Analysis of Effects				
Effect	F Value	Num DF	Den DF	Pr > F
_age_4g	24.41	3	328317	<.0001
_SEX	188.63	1	328319	<.0001
_HLTHPLN	28.28	1	328319	<.0001
_IMPRACE	29.34	5	328315	<.0001
_income_4g	83.06	3	328317	<.0001
_IMPRACE*_income_4g	7.51	15	328305	<.0001

5. Result reporting

- Write the method section to summarize the method used in data analysis.
- Prepare necessary tables and figures to illustrate the results. We often need a summary data, maybe a bivariate analysis table, and a table for the final model. Sometimes figures are better to illustrate the findings.
- Write the result section to summarize the key findings.
- Write the conclusion and limitations.

Statistical analysis

Data were summarized in mean (standard deviation) or frequency (percent). Complex samples procedures, proc surveymeans and proc surveyreg, in SAS 9.4 (SAS Institute Inc., Cary, NC) were used to compare the mean number of days physical health not good among the racial/ethnic populations with a Taylor series variance estimation method. When the number of days physical health not good was dichotomized into 0 or 1 or more days, SAS proc surveylogistic was used to compare the odds of having 1 or more days physical health not good among the racial/ethnic populations. In all analyses, the non-Hispanic White population was used as a reference group. Age, sex, and health insurance coverage were considered covariates, while income level was considered an effect modifier for race/ethnicity. A significance level of 0.05 was used.

Table 1 - Summary

Variable	Category	Weighted n	Mean or %	Std Err
# of days physical health not good		241,109,851	3.54	0.03
1 or more days physical health of good		77,426,077	32.11	0.16
Age	18-34	72,481,409	29.46	0.16
	35-54	79,156,834	32.17	0.16
	55-64	40,088,457	16.29	0.12
	65+	54,314,940	22.08	0.12
Sex	Male	119,902,241	48.73	0.17
	Female	126,139,399	51.27	0.17
Health insurance coverage	Have some form of insurance	214,432,112	91.41	0.11
	Do not have some form of health insurance	20,162,490	8.59	0.11
Race/ethnicity	White, Non-Hispanic	152,731,715	62.08	0.17
	Black, Non-Hispanic	28,639,654	11.64	0.11
	Asian, Non-Hispanic	14,461,978	5.88	0.12
	American Indian/Alaskan Native, Non-Hispanic	2,350,504	0.96	0.02
	Hispanic	42,298,969	17.19	0.15
	Other race, Non-Hispanic	5,558,820	2.26	0.04
Income level	\$100,000 or more	51,575,592	26.96	0.17
	\$50,000-\$100,000	55,555,176	29.04	0.17
	\$15,000-\$50,000	69,358,743	36.26	0.18
	Less than \$15,000	14,802,143	7.74	0.11

Table 2 - bivariate

Race/ethnicity	# of days physical health not good		1 or more days physical health of good	
	Mean	Std Err	%	Std Err
White, Non-Hispanic	3.64	0.03	33.34	0.17
Black, Non-Hispanic	3.77	0.08	33.15	0.48
Asian, Non-Hispanic	1.92	0.15	23.76	0.98
American Indian/Alaskan Native, Non-Hispanic	5.31	0.25	38.97	1.20
Hispanic	3.41	0.08	29.17	0.47
Other race, Non-Hispanic	4.05	0.15	34.32	0.83

Table 3 - Models

Outcome	Race/ethnicity ³	All	<\$15K	Income level		
				\$15K- \$50K	\$50K- \$100K	>\$100K
# of days physical health not good (mean difference) ¹	American Indian/Alaskan Native, Non-Hispanic	1.98	-0.73	0.90	0.20	0.94
	Asian, Non-Hispanic	-1.18	-4.89	-2.24	-0.58	-0.53
	Black, Non-Hispanic	0.34	-2.22	-0.74	-0.39	0.18
	Hispanic	0.29	-2.96	-1.13	0.08	0.19
	Other race, Non-Hispanic	0.81	-0.03	0.23	0.75	-0.16
1 or more days physical health of good (odds ratio) ²	American Indian/Alaskan Native, Non-Hispanic	1.34	0.96	1.07	0.93	1.09
	Asian, Non-Hispanic	0.66	0.36	0.52	0.78	0.82
	Black, Non-Hispanic	1.00	0.82	0.80	0.84	0.92
	Hispanic	0.87	0.47	0.59	0.93	1.04
	Other race, Non-Hispanic	1.08	0.77	0.95	1.02	0.93

1: linear regressions; 2: logistic regressions; 1,2: Models were adjusted for age, sex, and health insurance coverage; 3: White, Non-Hispanic was the reference group.

Results

- There were disparities among racial/ethnic populations in the number of days physical health not good. Compared to the non-Hispanic White population, the non-Hispanic Asian population had better physical health, the American Indian/Alaskan Native population had worse physical health, and the non-Hispanic Black and Hispanic populations had similar physical health overall.
- Specifically for people with the lowest income level, all populations had better physical health than the non-Hispanic White population with the non-Hispanic Asian the best, Hispanic the second, and Black the third. However, this disparity gradually diminished as people's income increased.

Conclusion and limitations

- This study finds racial/ethnic disparities in the number of days physical health not good using BRFSS 2021 data. This disparity was the worst for people with the lowest income level and gradually diminished as people's income increased.
- The BRFSS 2021 is a large national sample survey, but Florida was excluded from the survey because it failed to generate enough sample data. This impairs the generalizability of the results of this study.
- The BRFSS 2021 survey had about 2% missing data on the outcome, 4% missing data on health insurance coverage, and *21.5% on income level*. This study used pair-wise deletion in the analyses, which may cause subpopulation analysis results to vary from the overall analysis.