Which groups are at highest risk for low uptake of flu vaccine?

Chris Okitondo and Irene Cavros

3/18/23

# 1. Summary/Abstract

\*\*To be completed (PART 5)

# 2. Introduction

### 2.0.1 General Background Information

\*\*To be completed (PART 4)

### 2.0.2 Questions/Hypotheses to be addressed

Using the aforementioned database, our research question of interest is as follows:

Which groups are at highest risk for low uptake of flu vaccine?

# 3. Methods

### 3.0.1 Description of data and data source

The data source used for this project is the National Health Interview Survey (NHIS.) It is a harmonized set of data beginning in 1963 with information on general health status, acute and chronic illness, functional limitations, access to care, insurance coverage, and health behaviors for the U.S. population. On average, the survey covers 100,000 persons in 45,000 households each year. PUMS NHIS currently contains over 17,000 integrated variables from 1963 to present.

As the outcome of interest, researchers focused on those who, in this household survey between the years of 2012 and 2021, answered no when they were asked whether they had ANY flu vaccine in the past 12 months.

Variables assessed are as follows:

1. Demographics: age, sex, race, ethnicity, veteran status

2. Socioeconomic status: education level, total combined family income, employment status, paid sick leave as part of job benefits, ratio of family income to poverty threshold

3. General health: perceived health status, categorical BMI

4. Conditions: ever told had asthma, ever told had cancer, ever told had coronary heart disease

5. Health behaviors: alcohol drinking status, cigarette smoking status

6. Access to care: has usual place for medical care, medical care delayed due to cost (past 12 months)

7. Health insurance: health insurance coverage status

8. Mental health: anxiety level, depression level

#### 3.0.1.1 Study population

The raw data for our period of interest directly extracted from the NHIS IPUMS survey initially had had 792,916 observations. After processing the raw data and removing all observations with missing data in any of our predictors of interest, 106980 individuals were included in our analysis. Table 1 below provides a summary of demographic characteristics of our study population.

\*\*To be completed (PART 4)

### 3.0.2 Statistical analysis

*Explain anything related to your statistical analyses.*

\*\*To be completed (PART 4)

# 4. Results

### 4.0.1 Exploratory/Descriptive analysis

\*\* Table 1: Demographic characteristics of the study population \*\*

### 4.0.2 Loading Table 1 from the Exploratory Analysis

result\_table1 <- readRDS(here("results", "table1\_demographics.rds"))  
result\_table1

|  | no | yes | Overall |
| --- | --- | --- | --- |
|  | (N=58425) | (N=48555) | (N=106980) |
| **factor(sex)** | | | |
| male | 29881 (51.1%) | 20902 (43.0%) | 50783 (47.5%) |
| female | 28544 (48.9%) | 27653 (57.0%) | 56197 (52.5%) |
| **Age** | | | |
| Mean (SD) | 45.0 (16.1) | 54.6 (17.8) | 49.3 (17.6) |
| Median [Min, Max] | 44.0 [18.0, 85.0] | 57.0 [18.0, 85.0] | 49.0 [18.0, 85.0] |
| **factor(race)** | | | |
| white | 46388 (79.4%) | 40081 (82.5%) | 86469 (80.8%) |
| black | 8356 (14.3%) | 5090 (10.5%) | 13446 (12.6%) |
| native american | 643 (1.1%) | 496 (1.0%) | 1139 (1.1%) |
| asian | 3038 (5.2%) | 2888 (5.9%) | 5926 (5.5%) |
| **factor(hispanic)** | | | |
| non-hispanic | 49333 (84.4%) | 43787 (90.2%) | 93120 (87.0%) |
| hispanic | 9092 (15.6%) | 4768 (9.8%) | 13860 (13.0%) |
| **factor(education)** | | | |
| less\_than\_hs | 6784 (11.6%) | 4528 (9.3%) | 11312 (10.6%) |
| high\_school | 15161 (25.9%) | 10228 (21.1%) | 25389 (23.7%) |
| some\_college | 19099 (32.7%) | 14086 (29.0%) | 33185 (31.0%) |
| bachelor | 11690 (20.0%) | 11268 (23.2%) | 22958 (21.5%) |
| graduate | 5691 (9.7%) | 8445 (17.4%) | 14136 (13.2%) |
| **factor(fam\_income)** | | | |
| < $50,000 | 29164 (49.9%) | 20515 (42.3%) | 49679 (46.4%) |
| $50,000-$99,999 | 17097 (29.3%) | 14222 (29.3%) | 31319 (29.3%) |
| $100,000+ | 12164 (20.8%) | 13818 (28.5%) | 25982 (24.3%) |
| **factor(employ\_status)** | | | |
| unemployed | 2484 (4.3%) | 931 (1.9%) | 3415 (3.2%) |
| employed | 42869 (73.4%) | 29153 (60.0%) | 72022 (67.3%) |
| retired | 13072 (22.4%) | 18471 (38.0%) | 31543 (29.5%) |
| **factor(paid\_sick)** | | | |
| no | 27943 (47.8%) | 16920 (34.8%) | 44863 (41.9%) |
| yes | 30482 (52.2%) | 31635 (65.2%) | 62117 (58.1%) |
| **factor(health\_ins\_status)** | | | |
| no | 9935 (17.0%) | 1869 (3.8%) | 11804 (11.0%) |
| yes | 48490 (83.0%) | 46686 (96.2%) | 95176 (89.0%) |
| **factor(med\_access)** | | | |
| no | 11046 (18.9%) | 2547 (5.2%) | 13593 (12.7%) |
| yes | 47379 (81.1%) | 46008 (94.8%) | 93387 (87.3%) |
| **factor(med\_care\_delayed\_cost)** | | | |
| no | 51052 (87.4%) | 45059 (92.8%) | 96111 (89.8%) |
| yes | 7373 (12.6%) | 3496 (7.2%) | 10869 (10.2%) |
| **factor(health\_status)** | | | |
| excellent | 16670 (28.5%) | 11618 (23.9%) | 28288 (26.4%) |
| very\_good | 20113 (34.4%) | 16656 (34.3%) | 36769 (34.4%) |
| good | 15133 (25.9%) | 13376 (27.5%) | 28509 (26.6%) |
| fair | 5199 (8.9%) | 5261 (10.8%) | 10460 (9.8%) |
| poor | 1310 (2.2%) | 1644 (3.4%) | 2954 (2.8%) |
| **factor(bmi)** | | | |
| underweight | 964 (1.7%) | 633 (1.3%) | 1597 (1.5%) |
| average | 19469 (33.3%) | 15396 (31.7%) | 34865 (32.6%) |
| overweight | 20276 (34.7%) | 17002 (35.0%) | 37278 (34.8%) |
| obese | 17716 (30.3%) | 15524 (32.0%) | 33240 (31.1%) |
| **factor(asthma)** | | | |
| no | 51449 (88.1%) | 41304 (85.1%) | 92753 (86.7%) |
| yes | 6976 (11.9%) | 7251 (14.9%) | 14227 (13.3%) |
| **factor(cancer)** | | | |
| no | 54623 (93.5%) | 41766 (86.0%) | 96389 (90.1%) |
| yes | 3802 (6.5%) | 6789 (14.0%) | 10591 (9.9%) |
| **factor(chd)** | | | |
| no | 56727 (97.1%) | 45045 (92.8%) | 101772 (95.1%) |
| yes | 1698 (2.9%) | 3510 (7.2%) | 5208 (4.9%) |
| **factor(alcohol\_status)** | | | |
| nondrinker | 9219 (15.8%) | 7708 (15.9%) | 16927 (15.8%) |
| former\_drinker | 7791 (13.3%) | 8159 (16.8%) | 15950 (14.9%) |
| drinker | 41415 (70.9%) | 32688 (67.3%) | 74103 (69.3%) |
| **factor(smoking\_status)** | | | |
| nonsmoker | 34750 (59.5%) | 29338 (60.4%) | 64088 (59.9%) |
| former\_smoker | 12027 (20.6%) | 13685 (28.2%) | 25712 (24.0%) |
| smoker | 11648 (19.9%) | 5532 (11.4%) | 17180 (16.1%) |

\*\*

### 4.0.3 Full analysis

##### 4.0.3.0.1 Model 1: Focusing on only demogrphics characteristics using the flu vaccine uptake as our outcome

result\_table2 <- readRDS(here("results", "demographics\_model.rds"))  
result\_table2

term estimate std.error statistic  
1 (Intercept) -2.412331604 0.0334164527 -72.1899366  
2 age 0.034347387 0.0003974164 86.4266897  
3 sexfemale 0.355237904 0.0130976192 27.1223264  
4 raceblack -0.240890761 0.0204033476 -11.8064332  
5 racenative american 0.231326583 0.0632861504 3.6552481  
6 raceasian 0.114989390 0.0284889803 4.0362761  
7 hispanichispanic -0.177889489 0.0209758170 -8.4806942  
8 educationhigh\_school 0.004245455 0.0249208828 0.1703573  
9 educationsome\_college 0.185041675 0.0244858985 7.5570710  
10 educationbachelor 0.424188493 0.0265939096 15.9505879  
11 educationgraduate 0.701967529 0.0294263747 23.8550462  
12 fam\_income$50,000-$99,999 0.132219857 0.0159985840 8.2644724  
13 fam\_income$100,000+ 0.353068166 0.0179804150 19.6362635  
 p.value  
1 0.000000e+00  
2 0.000000e+00  
3 5.370848e-162  
4 3.615769e-32  
5 2.569334e-04  
6 5.430631e-05  
7 2.238543e-17  
8 8.647291e-01  
9 4.122468e-14  
10 2.822273e-57  
11 8.976145e-126  
12 1.403139e-16  
13 7.577252e-86

### 4.0.4 Model 2: Using all our predictors of interest in the model with flu vaccine uptake as our outcome

result\_table3 <- readRDS(here("results", "full\_model.rds"))  
result\_table3

term estimate std.error statistic  
1 (Intercept) -2.239921505 0.0379425281 -59.034588  
2 age 0.030696805 0.0004333642 70.833735  
3 sexfemale 0.347985133 0.0133861311 25.995945  
4 raceblack -0.236115703 0.0206769405 -11.419277  
5 racenative american 0.250033086 0.0637233155 3.923730  
6 raceasian 0.132068451 0.0288590192 4.576332  
7 hispanichispanic -0.192490023 0.0212801177 -9.045534  
8 educationhigh\_school 0.004827906 0.0252706377 0.191048  
9 educationsome\_college 0.158552543 0.0249416764 6.356932  
10 educationbachelor 0.378764844 0.0272631661 13.892915  
11 educationgraduate 0.647484649 0.0301351349 21.486038  
12 fam\_income$50,000-$99,999 0.130690475 0.0162269769 8.053902  
13 fam\_income$100,000+ 0.348323226 0.0183432206 18.989208  
14 asthmayes 0.334923861 0.0193313991 17.325381  
15 canceryes 0.279409838 0.0230819284 12.105134  
16 chdyes 0.492680238 0.0324541351 15.180815  
17 smoking\_statusformer\_smoker 0.056881954 0.0165999208 3.426640  
18 smoking\_statussmoker -0.401446680 0.0199201482 -20.152796  
19 alcohol\_statusformer\_drinker 0.057943801 0.0243219525 2.382366  
20 alcohol\_statusdrinker -0.022146202 0.0195529294 -1.132628  
 p.value  
1 0.000000e+00  
2 0.000000e+00  
3 5.503490e-149  
4 3.350107e-30  
5 8.718844e-05  
6 4.731995e-06  
7 1.489357e-19  
8 8.484880e-01  
9 2.058229e-10  
10 6.993054e-44  
11 2.103115e-102  
12 8.019570e-16  
13 2.094648e-80  
14 3.026826e-67  
15 9.920715e-34  
16 4.738537e-52  
17 6.110985e-04  
18 2.543362e-90  
19 1.720178e-02  
20 2.573704e-01

\*\*

# 5. Discussion

### 5.0.1 Summary and Interpretation

*Summarize what you did, what you found and what it means.*

\*\*To be completed (PART 5)

### 5.0.2 Strengths and Limitations

*Discuss what you perceive as strengths and limitations of your analysis.*

\*\*To be completed (PART 5)

### 5.0.3 Conclusions

*What are the main take-home messages?*

\*\*To be completed (PART 5)

# 6. References

*Include citations in your Rmd file using bibtex, the list of references will automatically be placed at the end*

\*\*To be completed (PART 5)