Project Milestone 2

Steve Chen

March 27, 2019

Basic Information

- Project Title: Unstable Housing and Hypertension-Related Emergency Department Hospitalizations among Health Center Patients
- Name: Steve Chen
- Email Address: steve.chen@emory.edu

Key Links

- GitHub repository: https://github.com/chenssteve/NRSG741final.git
- Data Source:
    - Health Center Patient Survey https://bphc.hrsa.gov/datareporting/research/hcpsurvey/index.html

Project Objectives

What is the primary focal question that you are trying to answer? What would you like to learn and accomplish?

The objective of this project is to determine whether there is an association between unstable housing and hypertension-associated emergency department utilization among patients seeking care at federally qualified health centers (FQHCs). The results of the study will inform the role of housing in determining health care utilization among safety net clinic patients.

Note that this question is completely different than the original study question proposed in Project Milestone 1. This is due to changing personal research interests as well as a desire to align the project's research outcomes with my MSPH thesis, which will leverage this dataset as well.

Data

From where and how are you acquiring your data? Provide a link to your data source.

The Health Center Patient Survey (HCPS) is a nationally representative survey of patients who receive care at federally-funded safety net health centers. The survey is sponosred by the Health Resources & Services Administration and provides data on how well health centers meet the health care needs of the medically underserved. The HCPS is administered every five years, with the most recent dataset obtained in 2014. Survey data is obtained from in-person, one-on-one interview with a nationally representative sample of health center patients.

The dataset is publicly available on the HRSA website for download. It is availale in SPSS, SAS, and STATA file formats.

Methods

The dataset was appropriately imported into R from the STATA file format.

```
library(foreign)
data <- read.dta("/Users/Steve/Box Sync/Emory MSPH/Spring 2019/NRSG 741/NRSG741final/PUF.dta")
```

Relevant variables were selected from the dataset.

- Identification and survey weights: zrid, analwt, vestr
- Hypertension indicators: con10, con10c
- Emergency room use due to hypertension: foll_conf3e
- Unstable housing indicators: liv6, liv7_r, liv4, liv1_r
- Sociodemographic covariates: intage_r, edit_gen, final_race, ins2, ins4, ins5, ins6, ins7, ins8, ins9, education, fpl, int4, dmo1, urban
- Facility indicators: ptype

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
subsetdata <- data %>% select("zrid", "analwt", "vestr", "verep", "con10", "con10c",
    "liv6", "liv7_r", "liv4", "liv1_r", "foll_conf3e", "intage_r", "edit_gen",
    "final_race", "ins2", "ins4", "ins5", "ins6", "ins7", "ins8", "ins9", "education",
    "fpl", "int4", "dmo1", "urban", "ptype")
```

The study sample included all adults (aged 18 years or older) who reported being told by a doctor or health professional that they had high blood pressure. The data was cleaned appropriately.

- Hypertension: con10/con10c: "YES"
- Age: intage_r: 18+

```
subsetdata1 <- subsetdata %>% filter(intage_r != "0-11" & intage_r != "12-17") %>%
    filter(con10 == "YES" | con10c == "YES")
```

Next, we created a new variable to identify those individuals who lived in unstable housing. Unstable housing was defined using similar critiera found in previous studies. They included those who reported that were staying in a house or apartment that they did not own or rent, those who did not have enough money to pay their rent or mortgage, and those who reported two or more moves within the past year. Those who reported none of these issues were classifed as having stable housing. Individuals experiencing homelessness were removed from the study sample, given their unique health care challenges faced that may be different from those faced by individuals living in unstable housing.

```r
# Remove individuals experiencing homelessness.
subsetdata2 <- subsetdata1 %>% filter(liv1_r == "A HOUSE, TOWNHOUSE, OR MOBILE HOME" |
    liv1_r == "AN APARTMENT OR CONDO")

# Create indicator variable for unstable housing.
subsetdata2$unstable <- NA

subsetdata2$unstable[subsetdata2$liv4 == "YES"] <- "NO"
subsetdata2$unstable[subsetdata2$liv6 == "NO"] <- "NO"
subsetdata2$unstable[subsetdata2$liv7_r == "0" | subsetdata2$liv7_r == "1"] <- "NO"
subsetdata2$unstable[subsetdata2$liv4 == "NO"] <- "YES"
subsetdata2$unstable[subsetdata2$liv6 == "YES"] <- "YES"
subsetdata2$unstable[subsetdata2$liv7_r == "2" | subsetdata2$liv7_r == "3" |
    subsetdata2$liv7_r == "4"] <- "YES"

# Remove individuals who did not answer (1) whether they owned or rented
# their home, (2) whether they had challenges affording rent/mortgage, or
# (3) the number of times they have moved.
subsetdata3 <- subsetdata2 %>% filter(!is.na(unstable))

# Confirm the dataset was cleaned.
library(janitor)
subsetdata3 %>% janitor::tabyl(liv4, unstable)
```

```
##        liv4   NO YES
##      Refusal   0   0
##   Don't know   0   0
##         YES 1327 582
##          NO    0 232
```

```r
subsetdata3 %>% janitor::tabyl(liv6, unstable)
```

```
## Warning: Factor `liv6` contains implicit NA, consider using
## `forcats::fct_explicit_na`
```

```
##        liv6   NO YES
##      Refusal   1   0
##   Don't know   1   1
##         YES    0 493
##          NO 1325  88
##        <NA>    0 232
```

```r
subsetdata3 %>% janitor::tabyl(liv7_r, unstable)
```

```
##  liv7_r   NO YES
##       0 1122 358
##       1  205  71
##       2    0  92
##       3    0  32
##       4    0  29
##      NA    0 232
```

We then removed and cleaned the focal dependent variable measuring self-reported hypertension-related ER utilization. As part of the survey, participants were asked whether they visited the hospital or emergency room due to hypertension. We are only interested in those who responded "Yes" or "No."

```
subsetdata4 <- subsetdata3
subsetdata4$foll_conf3e[subsetdata4$foll_conf3e == "Don't know"] <- NA

subsetdata4 <- subsetdata4 %>% filter(!is.na(foll_conf3e))

subsetdata4 %>% janitor::tabyl(foll_conf3e, unstable)
```

```
##  foll_conf3e   NO YES
##      Refusal    0   0
##   Don't know    0   0
##          YES  100 102
##           NO 1061 587
```

```
finalsub <- subsetdata4

finalsub$er <- relevel(finalsub$foll_conf3e, ref = "NO")
```

Lastly, we included and cleaned the sociodemographic covariate measures, including age, gender, race, education, income, primary language, nationality, urbanicity, insurance, and center type.

```
# Create indicator variable for insured / uninsred.
finalsub$ins <- NA

finalsub$ins[finalsub$ins2 == "NO"] <- "NO"
finalsub$ins[finalsub$ins4 == "NO"] <- "NO"
finalsub$ins[finalsub$ins5 == "NO"] <- "NO"
finalsub$ins[finalsub$ins6 == "NO"] <- "NO"
finalsub$ins[finalsub$ins7 == "NO"] <- "NO"
finalsub$ins[finalsub$ins8 == "NO"] <- "NO"
finalsub$ins[finalsub$ins9 == "NO"] <- "NO"

finalsub$ins[finalsub$ins2 == "YES"] <- "YES"
finalsub$ins[finalsub$ins4 == "YES"] <- "YES"
finalsub$ins[finalsub$ins5 == "YES"] <- "YES"
finalsub$ins[finalsub$ins6 == "YES"] <- "YES"
finalsub$ins[finalsub$ins7 == "YES"] <- "YES"
finalsub$ins[finalsub$ins8 == "YES"] <- "YES"
finalsub$ins[finalsub$ins9 == "YES"] <- "YES"

finalsub$ins.f <- factor(finalsub$ins, levels = c("YES", "NO"))

# Clean missing values for race and nationality.
finalsub$race_char <- as.character(finalsub$final_race)
finalsub$race_char[finalsub$race_char == "Refusal"] <- NA
finalsub$race <- factor(finalsub$race_char, levels = c("NonHispanic White",
    "NonHispanic Black", "NonHispanic Asian", "NonHispanic Other", "Hispanic"))

finalsub$dmo1[finalsub$dmo1 == "Don't know"] <- NA
```

```r
# Convert education into a factor to accurately present it as a categorical
# variable.
finalsub$education.f <- factor(finalsub$education, levels = 1:3, labels = c("Less than high school",
    "High school", "More than high school"))

# Rename variables for easy presentation.
finalsub$ins <- finalsub$ins.f
finalsub$age <- finalsub$intage_r
finalsub$gender <- finalsub$edit_gen
finalsub$education <- finalsub$education.f
finalsub$fpl <- finalsub$fpl
finalsub$lang <- finalsub$int4
finalsub$nation <- finalsub$dmo1
finalsub$urban <- finalsub$urban
finalsub$ptype <- finalsub$ptype

finalsub_demog <- finalsub %>% select("unstable", "er", "gender", "race", "age",
    "education", "fpl", "lang", "nation", "urban", "ins", "ptype", "zrid", "vestr",
    "analwt")

# Drop missing categories
finalsub_demog <- droplevels(finalsub_demog)
```

Analysis

We first conducted descriptive analyses of both unweighted and weighted data. Statistical comparisons for significance were conducted using chi square and ANOVA tests.

```r
library(tableone)
vars = c("age", "gender", "race", "education", "fpl", "lang", "nation", "urban",
    "ins", "ptype")

# Unweighted
tab1_uw <- tableone::CreateTableOne(vars = vars, strata = "unstable", data = finalsub_demog)

tab1_uw
```

```
##                          Stratified by unstable
##                           NO           YES          p        test
##   n                       1161         689
##   age (%)                                           <0.001
##      18-20                  2 ( 0.2)     1 ( 0.1)
##      21-25                  9 ( 0.8)     7 ( 1.0)
##      26-34                 35 ( 3.0)    33 ( 4.8)
##      35-44                117 (10.1)   112 (16.3)
##      45-54                286 (24.6)   238 (34.5)
##      55-64                426 (36.7)   236 (34.3)
##      65-74                211 (18.2)    50 ( 7.3)
##      75 or older           75 ( 6.5)    12 ( 1.7)
##   gender = Female (%)     729 (62.8)   439 (63.7)   0.727
##   race (%)                                          <0.001
##      NonHispanic White    293 (25.3)   195 (28.3)
```

```
##      NonHispanic Black              329 (28.4)  204 (29.6)
##      NonHispanic Asian               85 ( 7.3)   17 ( 2.5)
##      NonHispanic Other              123 (10.6)   52 ( 7.5)
##      Hispanic                       329 (28.4)  221 (32.1)
##   education (%)                                                 0.020
##      Less than high school          526 (45.4)  276 (40.1)
##      High school                    306 (26.4)  177 (25.7)
##      More than high school          327 (28.2)  235 (34.2)
##   fpl (%)                                                      <0.001
##      Less than or Equal to 100% FPL 659 (57.1)  479 (69.6)
##      101% to 138% FPL               214 (18.5)  112 (16.3)
##      139% to 199% FPL               128 (11.1)   67 ( 9.7)
##      200% to 299% FPL                85 ( 7.4)   18 ( 2.6)
##      300% to 399% FPL                30 ( 2.6)    8 ( 1.2)
##      400% or More than FPL           38 ( 3.3)    4 ( 0.6)
##   lang = NO (%)                      741 (63.8)  474 (68.8)  0.033
##   nation = NO (%)                    341 (29.4)  187 (27.3)  0.359
##   urban = Rural (%)                  397 (34.2)  207 (30.0)  0.074
##   ins = NO (%)                        26 ( 2.2)   18 ( 2.6)  0.725
##   ptype (%)                                                  <0.001
##      Public Housing Primary Care    123 (10.6)   55 ( 8.0)
##      Migrant Health Center          102 ( 8.8)   78 (11.3)
##      Health Care for the Homeless    42 ( 3.6)  150 (21.8)
##      Community Health Center        894 (77.0)  406 (58.9)
```

```
# summary(tab1_uw)

# Weighted
library(survey)
```

```
## Loading required package: grid

## Loading required package: Matrix

## Loading required package: survival

##
## Attaching package: 'survey'

## The following object is masked from 'package:graphics':
##
##     dotchart
```

```
finalsub_demog$analwt <- as.numeric(finalsub_demog$analwt)
hcpssurvey <- svydesign(ids = ~zrid, strata = ~vestr, weights = ~analwt, data = finalsub_demog)

tab1 <- tableone::svyCreateTableOne(vars = vars, strata = "unstable", data = hcpssurvey)
tab1
```

```
##                                 Stratified by unstable
##                                  NO            YES
##   n                             3711604.5     2185213.1
```

```
##   age (%)
##      18-20                                37753.0 ( 1.0)    51597.3 ( 2.4)
##      21-25                                10280.5 ( 0.3)    51378.1 ( 2.4)
##      26-34                               246296.0 ( 6.6)   310317.1 (14.2)
##      35-44                               574714.6 (15.5)   410436.7 (18.8)
##      45-54                               847460.7 (22.8)   510616.1 (23.4)
##      55-64                              1024534.5 (27.6)   647969.1 (29.7)
##      65-74                               652665.9 (17.6)   142542.2 ( 6.5)
##      75 or older                         317899.3 ( 8.6)    60356.7 ( 2.8)
##   gender = Female (%)                    1836359.7 (49.5)  1326247.9 (60.7)
##   race (%)
##      NonHispanic White                  2212733.5 (59.6)  1387091.7 (63.5)
##      NonHispanic Black                   705509.4 (19.0)   414861.3 (19.0)
##      NonHispanic Asian                    52476.4 ( 1.4)    11611.1 ( 0.5)
##      NonHispanic Other                   203865.0 ( 5.5)    70375.2 ( 3.2)
##      Hispanic                            535145.4 (14.4)   301273.8 (13.8)
##   education (%)
##      Less than high school              1395031.7 (37.8)   732949.8 (33.6)
##      High school                         932465.4 (25.3)   537220.9 (24.6)
##      More than high school              1361862.2 (36.9)   913467.3 (41.8)
##   fpl (%)
##      Less than or Equal to 100% FPL 1903184.0 (52.3)  1334612.6 (61.3)
##      101% to 138% FPL                    557804.6 (15.3)   357676.9 (16.4)
##      139% to 199% FPL                    382536.2 (10.5)   320284.2 (14.7)
##      200% to 299% FPL                    467506.1 (12.9)    20328.6 ( 0.9)
##      300% to 399% FPL                    131277.9 ( 3.6)   120879.4 ( 5.6)
##      400% or More than FPL               194037.5 ( 5.3)    22984.8 ( 1.1)
##   lang = NO (%)                          2993506.5 (80.7)  1894588.7 (86.7)
##   nation = NO (%)                         513631.7 (13.8)   253696.8 (11.6)
##   urban = Rural (%)                      2101371.2 (56.6)  1325680.8 (60.7)
##   ins = NO (%)                             65807.8 ( 1.8)    15631.2 ( 0.7)
##   ptype (%)
##      Public Housing Primary Care          37663.9 ( 1.0)    19119.5 ( 0.9)
##      Migrant Health Center                63510.8 ( 1.7)    37900.8 ( 1.7)
##      Health Care for the Homeless        112788.7 ( 3.0)    96641.3 ( 4.4)
##      Community Health Center            3497641.1 (94.2)  2031551.5 (93.0)
##                                         Stratified by unstable
##                                         p       test
##   n
##   age (%)                                0.012
##      18-20
##      21-25
##      26-34
##      35-44
##      45-54
##      55-64
##      65-74
##      75 or older
##   gender = Female (%)                    0.060
##   race (%)                               0.586
##      NonHispanic White
##      NonHispanic Black
##      NonHispanic Asian
##      NonHispanic Other
```

```
##        Hispanic
##    education (%)                              0.665
##        Less than high school
##        High school
##        More than high school
##    fpl (%)                                    0.002
##        Less than or Equal to 100% FPL
##        101% to 138% FPL
##        139% to 199% FPL
##        200% to 299% FPL
##        300% to 399% FPL
##        400% or More than FPL
##    lang = NO (%)                              0.081
##    nation = NO (%)                            0.459
##    urban = Rural (%)                          0.472
##    ins = NO (%)                               0.063
##    ptype (%)                                  0.545
##        Public Housing Primary Care
##        Migrant Health Center
##        Health Care for the Homeless
##        Community Health Center
```

```
# summary(tab1)
```

We then used logistic regression to analyze the association between unstable housing and hypertension-related emergency department visits, while controlling for the various sociodemographic covariates. Three models were created, two reduced models using unweighted and weighted data, and one comprehensive model including sociodemographic covariates using weighted data.

```
# Basic model, unadjusted for survey weights
model_basic_unadjust <- glm(er ~ unstable, family = binomial, data = finalsub_demog)
summary(model_basic_unadjust)
```

```
##
## Call:
## glm(formula = er ~ unstable, family = binomial, data = finalsub_demog)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.5661  -0.5661  -0.4244  -0.4244   2.2144
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.3618     0.1046 -22.578  < 2e-16 ***
## unstableYES    0.6117     0.1498   4.083 4.45e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1275.8  on 1849  degrees of freedom
## Residual deviance: 1259.3  on 1848  degrees of freedom
## AIC: 1263.3
```

```
##
## Number of Fisher Scoring iterations: 5
```

```
exp(coef(model_basic_unadjust))
```

```
## (Intercept) unstableYES
##  0.09425071  1.84364566
```

```
# Basic model, adjusted for survey weights
model_basic <- svyglm(er ~ unstable, family = binomial(link = "logit"), design = hcpssurvey)
```

```
## Warning in eval(family$initialize): non-integer #successes in a binomial
## glm!
```

```
summary(model_basic)
```

```
##
## Call:
## svyglm(formula = er ~ unstable, design = hcpssurvey, family = binomial(link = "logit"))
##
## Survey design:
## svydesign(ids = ~zrid, strata = ~vestr, weights = ~analwt, data = finalsub_demog)
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -2.5678     0.2314  -11.10   <2e-16 ***
## unstableYES   0.6178     0.3471    1.78   0.0753 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1.000541)
##
## Number of Fisher Scoring iterations: 4
```

```
exp(coef(model_basic))
```

```
## (Intercept) unstableYES
##  0.07670666  1.85478941
```

```
# Detailed model, adjusted for survey weights
model <- svyglm(er ~ unstable + age + gender + race + education + fpl + lang +
    nation + urban + ins + ptype, family = binomial(link = "logit"), design = hcpssurvey)
```

```
## Warning in eval(family$initialize): non-integer #successes in a binomial
## glm!
```

```
summary(model)
```

```
##
## Call:
## svyglm(formula = er ~ unstable + age + gender + race + education +
##     fpl + lang + nation + urban + ins + ptype, design = hcpssurvey,
##     family = binomial(link = "logit"))
##
## Survey design:
## svydesign(ids = ~zrid, strata = ~vestr, weights = ~analwt, data = finalsub_demog)
##
## Coefficients:
##                                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)                     -16.65893    1.33807 -12.450  < 2e-16
## unstableYES                       0.37839    0.39647   0.954   0.3400
## age21-25                         14.74390    1.50302   9.810  < 2e-16
## age26-34                         13.72258    1.20828  11.357  < 2e-16
## age35-44                         12.72610    1.22837  10.360  < 2e-16
## age45-54                         13.24186    1.12074  11.815  < 2e-16
## age55-64                         12.89239    1.24019  10.396  < 2e-16
## age65-74                         12.82167    1.16724  10.985  < 2e-16
## age75 or older                   11.88090    1.21795   9.755  < 2e-16
## genderFemale                     -0.16614    0.32610  -0.509   0.6105
## raceNonHispanic Black             1.70348    0.43300   3.934 8.67e-05
## raceNonHispanic Asian             0.26479    0.96342   0.275   0.7835
## raceNonHispanic Other             2.13533    0.66286   3.221   0.0013
## raceHispanic                      0.68701    0.57015   1.205   0.2284
## educationHigh school             -0.54072    0.47487  -1.139   0.2550
## educationMore than high school    0.50327    0.39364   1.279   0.2012
## fpl101% to 138% FPL              -0.95468    0.59836  -1.595   0.1108
## fpl139% to 199% FPL              -0.58263    0.61668  -0.945   0.3449
## fpl200% to 299% FPL              -1.43034    0.78341  -1.826   0.0680
## fpl300% to 399% FPL              -2.97599    0.96166  -3.095   0.0020
## fpl400% or More than FPL         -1.43079    1.20184  -1.191   0.2340
## langNO                            0.02361    0.63283   0.037   0.9702
## nationNO                          0.56914    0.56638   1.005   0.3151
## urbanRural                        0.40476    0.34065   1.188   0.2349
## insNO                            -1.07013    0.80885  -1.323   0.1860
## ptypeMigrant Health Center        0.13208    0.62945   0.210   0.8338
## ptypeHealth Care for the Homeless -0.27049    0.74131  -0.365   0.7152
## ptypeCommunity Health Center      0.40438    0.43325   0.933   0.3508
##
## (Intercept)                     ***
## unstableYES
## age21-25                        ***
## age26-34                        ***
## age35-44                        ***
## age45-54                        ***
## age55-64                        ***
## age65-74                        ***
## age75 or older                  ***
## genderFemale
## raceNonHispanic Black           ***
## raceNonHispanic Asian
## raceNonHispanic Other           **
## raceHispanic
```

```
## educationHigh school
## educationMore than high school
## fpl101% to 138% FPL
## fpl139% to 199% FPL
## fpl200% to 299% FPL                          .
## fpl300% to 399% FPL                         **
## fpl400% or More than FPL
## langNO
## nationNO
## urbanRural
## insNO
## ptypeMigrant Health Center
## ptypeHealth Care for the Homeless
## ptypeCommunity Health Center
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1.183402)
##
## Number of Fisher Scoring iterations: 14
```

exp(coef(model))

```
##                      (Intercept)                        unstableYES
##                     5.822649e-08                       1.459936e+00
##                           age21-25                           age26-34
##                     2.530436e+06                       9.112607e+05
##                           age35-44                           age45-54
##                     3.364148e+05                       5.634637e+05
##                           age55-64                           age65-74
##                     3.972776e+05                       3.701517e+05
##                      age75 or older                       genderFemale
##                     1.444809e+05                       8.469298e-01
##            raceNonHispanic Black               raceNonHispanic Asian
##                     5.493003e+00                       1.303152e+00
##            raceNonHispanic Other                        raceHispanic
##                     8.459801e+00                       1.987767e+00
##            educationHigh school     educationMore than high school
##                     5.823290e-01                       1.654119e+00
##              fpl101% to 138% FPL               fpl139% to 199% FPL
##                     3.849366e-01                       5.584273e-01
##              fpl200% to 299% FPL               fpl300% to 399% FPL
##                     2.392278e-01                       5.099683e-02
##         fpl400% or More than FPL                             langNO
##                     2.391194e-01                       1.023894e+00
##                          nationNO                          urbanRural
##                     1.766744e+00                       1.498945e+00
##                            insNO       ptypeMigrant Health Center
##                     3.429636e-01                       1.141194e+00
## ptypeHealth Care for the Homeless     ptypeCommunity Health Center
##                     7.630042e-01                       1.498369e+00
```

Next Steps

The next steps will involve writing up study analysis results and conclusions. Careful consideration will be made as to how best to visually present this data.