

# p8123\_final

Farizah Rob

2023-12-14

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.4      v readr      2.1.4
## v forcats    1.0.0      v stringr   1.5.1
## v ggplot2    3.4.4      v tibble    3.2.1
## v lubridate  1.9.3      v tidyr     1.3.0
## v purrr      1.0.2
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(survey)
```

```
## Loading required package: grid
## Loading required package: Matrix
##
## Attaching package: 'Matrix'
##
## The following objects are masked from 'package:tidyr':
##
##   expand, pack, unpack
##
## Loading required package: survival
##
## Attaching package: 'survey'
##
## The following object is masked from 'package:graphics':
##
##   dotchart
```

```
library(ggsurvey)
```

```
## Loading required package: hexbin
```

```
library(tableone)
library(gtsummary)
```

```
load("./data/nhanes_data.rda") #put files into data folder so everyone can run this
load("./data/nhanes_key.rda")
load("./data/key_guide.rda")
```

**Problem statement:** Controlling blood pressure (BP) reduces the risk for cardiovascular disease. However, the prevalence of BP control (i.e., systolic BP < 140 and diastolic BP < 90) among US adults with hypertension has decreased since 2013. We invite teams to analyze publicly available data from US adults to help identify potential causes or correlates of worsening BP control among US adults with hypertension over the past decade, as this may allow for development of effective interventions to help control BP and prevent cardiovascular disease.

**Svydesign and subset for hypertension patients 2013 onwards**

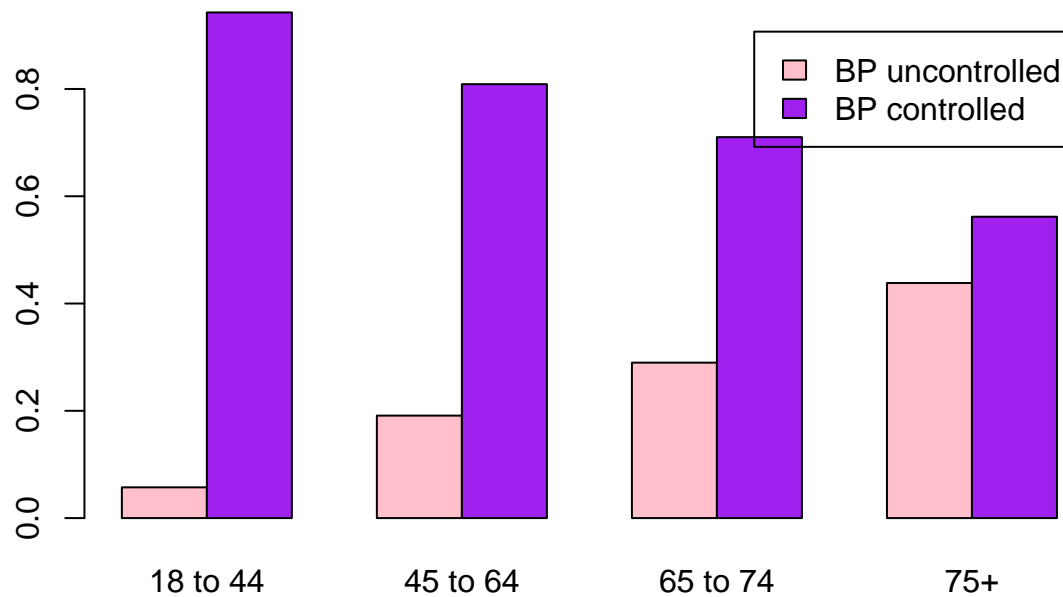
```
dstrat <- svydesign(ids=~svy_psu, strata=~svy_strata, weights=~svy_weight_mec, data=nhanes_data, nest=TRUE)
htsub <- subset(dstrat, svy_subpop_htn==1) #limited to hypertension patients
htsub2 <- subset(htsub, svy_year=="2013-2014" | svy_year=="2015-2016" | svy_year=="2017-2020") #limited to 2013 onwards
```

**Some EDA**

The **outcome** of interest -> **bp\_control\_140\_90** OR **bp\_uncontrolled\_140\_90**

```
#age -> so for older people, BP control is definitely lower
age <- svyby(~factor(bp_control_140_90), ~factor(demo_age_cat), design=htsub2, svymean, na.rm=TRUE)
barplot(age, legend.text=c("BP uncontrolled", "BP controlled"), col=c("pink", "purple"), main="Proportion of BP control by age group")
```

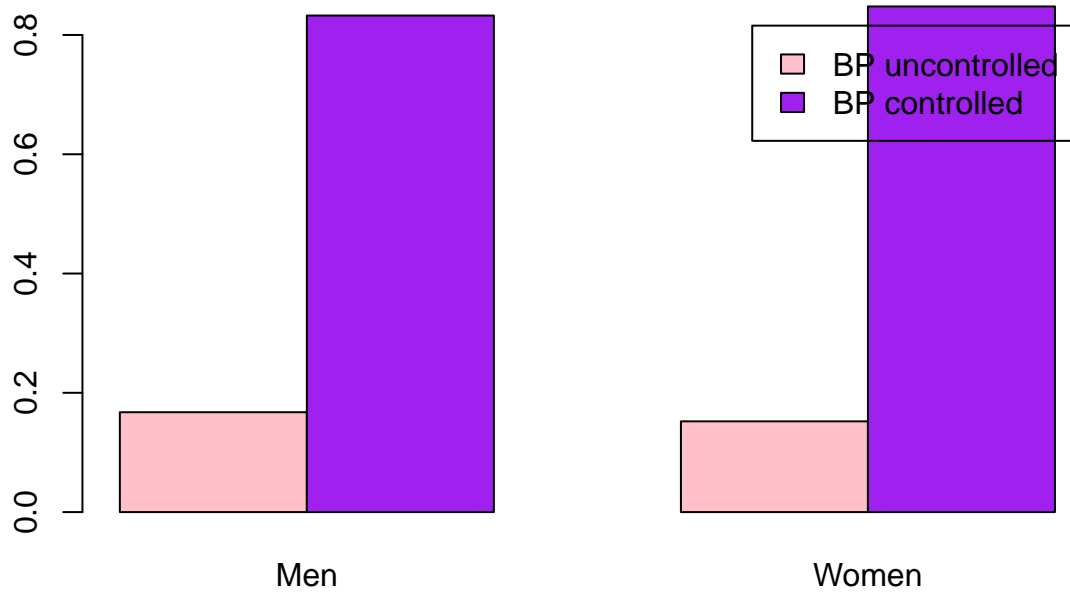
**Proportion of BP control by age category**



*#gender -> not really any difference*

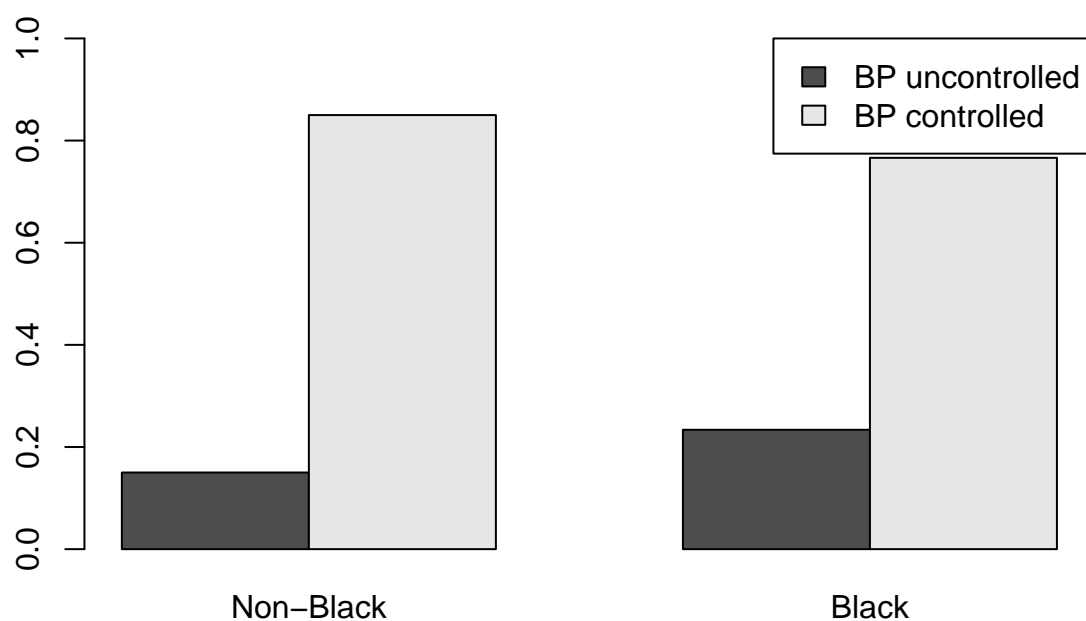
```
gender <- svyby(~factor(bp_control_140_90), ~factor(demo_gender), design=htsub2, svymean, na.rm=TRUE)
barplot(gender, legend.text=c("BP uncontrolled", "BP controlled"), col=c("pink", "purple"), main="Proportion of BP control by age category")
```

## Proportion of BP control by gender



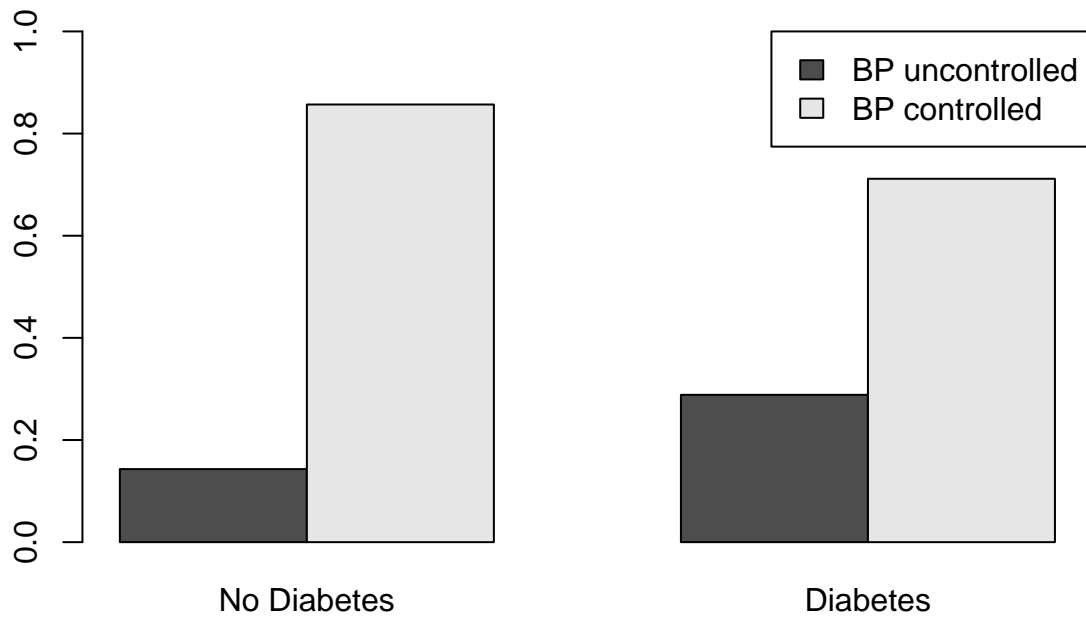
```
#race -> for Black people vs non-Black people, BP control is lower
race <- svyby(~factor(bp_control_140_90), ~factor(demo_race_black), design=htsub2, svymean, na.rm=TRUE)
barplot(race,
  legend.text=c("BP uncontrolled", "BP controlled"),
  names.arg=c("Non-Black", "Black"),
  ylim=c(0, 1.0),
  args.legend = list(x = "topright"),
  main="Proportion of BP control by race (non-Black vs Black)")
```

### Proportion of BP control by race (non-Black vs Black)



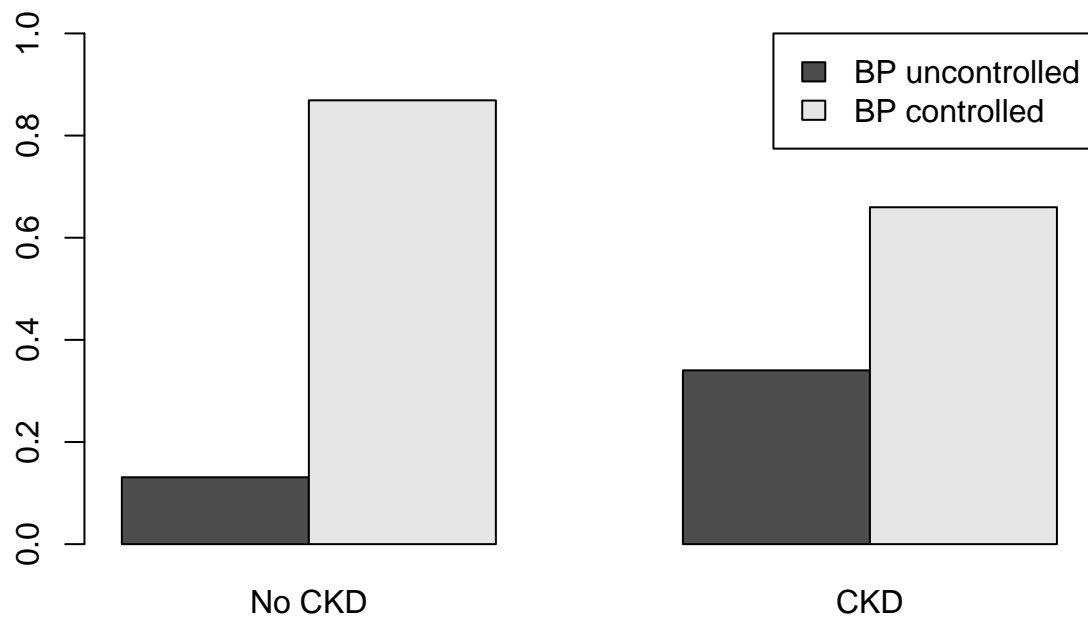
```
#diabetes
diabetes <- svyby(~factor(bp_control_140_90), ~factor(cc_diabetes), design=htsub2, svymean, na.rm=TRUE)
barplot(diabetes,
  legend.text=c("BP uncontrolled", "BP controlled"),
  names.arg=c("No Diabetes", "Diabetes"),
  ylim=c(0, 1.0),
  args.legend = list(x = "topright"),
  main="Proportion of BP control by diabetes")
```

### Proportion of BP control by diabetes



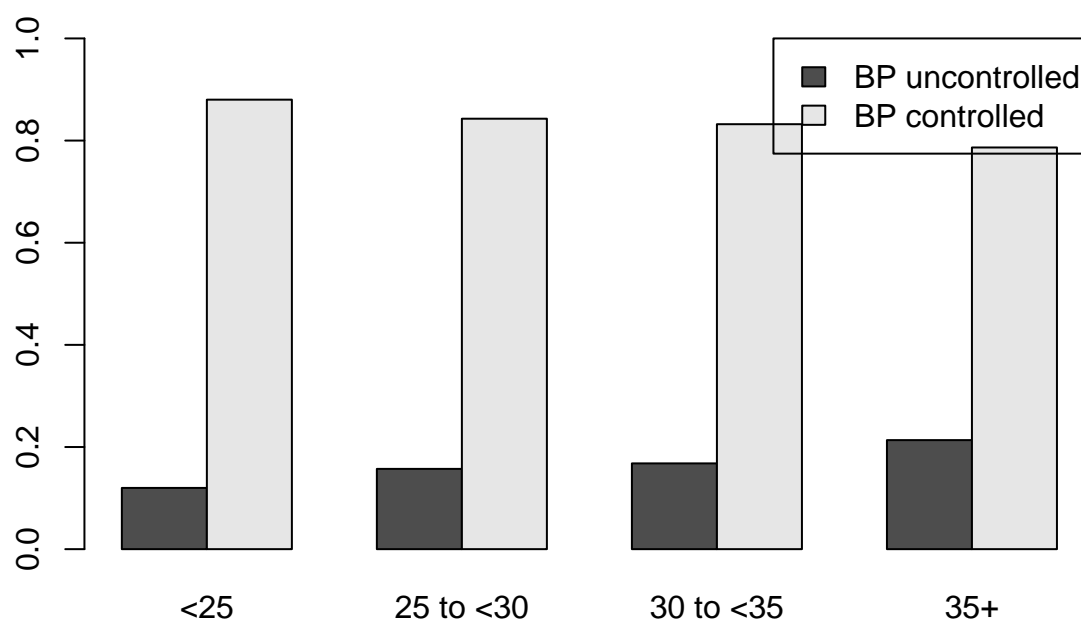
```
#ckd
ckd <- svyby(~factor(bp_control_140_90), ~factor(cc_ckd), design=htsub2, svymean, na.rm=TRUE)
barplot(ckd,
  legend.text=c("BP uncontrolled", "BP controlled"),
  names.arg=c("No CKD", "CKD"),
  ylim=c(0, 1.0),
  args.legend = list(x = "topright"),
  main="Proportion of BP control by CKD")
```

### Proportion of BP control by CKD



```
#bmi
bmi <- svyby(~factor(bp_control_140_90), ~factor(cc_bmi), design=htsub2, svymean, na.rm=TRUE)
barplot(bmi,
  legend.text=c("BP uncontrolled", "BP controlled"),
  ylim=c(0, 1.0),
  args.legend = list(x = "topright"),
  main="Proportion of BP control by BMI")
```

## Proportion of BP control by BMI



*#meds recommended -> 100% for those with uncontrolled BP*  
*# look at med recommended and meds taken ratio by different subgroups?*

```
table(nhanes_data$bp_med_recommended_jnc7, nhanes_data$bp_med_use) #48% of people recommended medication
```

```
##
##           No    Yes
## No  37755      0
## Yes  7062 14690
```

```
svyCreateCatTable(vars= c("bp_med_recommended_jnc7" ,"bp_med_use"),
                  strata="demo_race",
                  data=htsub2) #across all races, the % of med use is lower than % recommended -? this
```

```
##
##           Stratified by demo_race
##           Non-Hispanic White  Non-Hispanic Black
## n  443841370.8  78256419.3
## bp_med_recommended_jnc7 = Yes (%) 156173388.2 (35.2) 33534720.2 (42.9)
## bp_med_use = Yes (%) 111473163.2 (25.1) 23255215.8 (29.7)
##
##           Stratified by demo_race
##           Non-Hispanic Asian  Hispanic
## n  37820092.9  105372525.4
## bp_med_recommended_jnc7 = Yes (%) 11080831.4 (29.3) 26424308.7 (25.1)
## bp_med_use = Yes (%) 6822774.8 (18.0) 15828214.5 (15.0)
```



```
##                               Stratified by demo_race
##                               Other           p       test
##    n                          24220134.3
##    bp_med_recommended_jnc7 = Yes (%) 8029041.0 (33.2) <0.001
##    bp_med_use = Yes (%)             5443707.4 (22.5) <0.001

#for table 1 -> age category, gender, race, BMI, smoking status, diabetes, CKD, cholesterol, recommended
# removed "bp_med_recommended_jnc7", "bp_med_recommended_accaha" for space

# didn't they say something in class about the Rao and Scott being wrong in R? not sure if we should in

reset_gtsummary_theme()
theme_gtsummary_compact()
```

```
## Setting theme 'Compact'
```

```
tab1 <- tbl_svysummary(data = htsub2, by = "bp_control_140_90", include = c("demo_age_cat", "demo_gender",
  label = list(demo_age_cat ~ "Age",
    demo_gender ~ "Gender",
    demo_race ~ "Race",
    cc_bmi ~ "BMI",
    cc_ckd ~ "Chronic Kidney Disease",
    cc_diabetes ~ "Diabetes",
    cc_cvd_any ~ "Cardiovascular Disease",
    bp_med_use ~ "BP Medication Use",
    chol_total ~ "Total Cholesterol",
    chol_hdl ~ "HDL",
    chol_ldl ~ "LDL",
    chol_trig ~ "Triglycerides",
    chol_med_use ~ "Cholesterol Medication Use")) %>% add_p() %>% add_n()

  italicize_levels()

tab1
```

```
## Table printed with 'knitr::kable()', not {gt}. Learn why at
## https://www.danielsjoberg.com/gtsummary/articles/rmarkdown.html
## To suppress this message, include 'message = FALSE' in code chunk header.
```

Table 1: Subject Characteristics (Survey Weighted)

Characteristic	No, N = 110,019,543	Yes, N = 579,490,999	p-value
<b>Age, n (%)</b>			<0.001
18 to 44	18,218,707 (17%)	299,580,713 (52%)	
45 to 64	45,525,230 (41%)	192,746,908 (33%)	
65 to 74	23,816,079 (22%)	58,369,736 (10%)	
75+	22,459,527 (20%)	28,793,643 (5.0%)	
<b>Gender, n (%)</b>			0.050
Men	55,991,881 (51%)	278,426,528 (48%)	
Women	54,027,662 (49%)	301,064,472 (52%)	
<b>Race, n (%)</b>			<0.001

Characteristic	No, N = 110,019,543	Yes, N = 579,490,999	p-value
<i>Non-Hispanic White</i>	68,928,072 (63%)	374,913,299 (65%)	
<i>Non-Hispanic Black</i>	18,291,583 (17%)	59,964,836 (10%)	
<i>Non-Hispanic Asian</i>	5,715,421 (5.2%)	32,104,671 (5.5%)	
<i>Hispanic</i>	13,404,152 (12%)	91,968,373 (16%)	
<i>Other</i>	3,680,315 (3.3%)	20,539,820 (3.5%)	
<b>BMI, n (%)</b>			<0.001
<i>&lt;25</i>	23,597,988 (22%)	173,248,879 (30%)	
<i>25 to &lt;30</i>	33,526,572 (31%)	179,795,985 (32%)	
<i>30 to &lt;35</i>	23,810,864 (22%)	117,994,322 (21%)	
<i>35+</i>	26,600,214 (25%)	97,987,747 (17%)	
<i>Unknown</i>	2,483,906	10,464,066	
<b>Chronic Kidney Disease, n (%)</b>	32,045,575 (29%)	62,111,324 (11%)	<0.001
<b>Diabetes, n (%)</b>	22,552,554 (20%)	55,629,143 (9.6%)	<0.001
<b>Cardiovascular Disease, n (%)</b>	15,856,616 (14%)	40,856,909 (7.1%)	<0.001
<b>BP Medication Use, n (%)</b>	51,549,103 (47%)	111,273,973 (19%)	<0.001
<b>Total Cholesterol, Median (IQR)</b>	192 (163, 222)	183 (158, 211)	<0.001
<i>Unknown</i>	63,100,407	325,591,825	
<b>HDL, Median (IQR)</b>	52 (42, 65)	52 (43, 63)	>0.9
<i>Unknown</i>	63,100,407	325,591,825	
<b>LDL, Median (IQR)</b>	112 (88, 139)	109 (86, 133)	0.013
<i>Unknown</i>	63,221,679	325,731,647	
<b>Triglycerides, Median (IQR)</b>	105 (72, 155)	89 (61, 135)	<0.001
<i>Unknown</i>	63,221,679	325,731,647	
<b>Cholesterol Medication Use, n (%)</b>	16,391,227 (34%)	47,895,434 (19%)	<0.001
<i>Unknown</i>	62,505,218	323,610,642	

## Regression models - really don't know which variables to select

using the `bp_uncontrolled_140_90` variable instead of `bp_control` for regression

```
m1 <- svyglm(bp_uncontrolled_140_90 ~ demo_age_cat + demo_race + demo_gender + cc_smoke + cc_bmi + cc_diabetes + cc_ckd + cc_cvd_any + chol_trig + chol_ldl + chol_hdl, design = htsub2, family = quasibinomial())

#need to set BMI reference value to normal level rather than underweight

summary(m1)
```

```
##
## Call:
## svyglm(formula = bp_uncontrolled_140_90 ~ demo_age_cat + demo_race +
##   demo_gender + cc_smoke + cc_bmi + cc_diabetes + cc_ckd +
##   cc_cvd_any + chol_trig + chol_ldl + chol_hdl, design = htsub2,
##   family = quasibinomial())
##
## Survey design:
## subset(htsub, svy_year == "2013-2014" | svy_year == "2015-2016" |
##   svy_year == "2017-2020")
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -4.2784289   0.2847878  -15.023   < 2e-16 ***
```

```
## demo_age_cat45 to 64      1.1528233  0.1252890   9.201 5.47e-11 ***
## demo_age_cat65 to 74      1.7627753  0.1598648  11.027 4.28e-13 ***
## demo_age_cat75+          2.4495017  0.1681808  14.565 < 2e-16 ***
## demo_raceNon-Hispanic Black 0.8670093  0.1173396   7.389 1.02e-08 ***
## demo_raceNon-Hispanic Asian 0.3999404  0.1137010   3.517 0.001199 **
## demo_raceHispanic         0.1559047  0.1091321   1.429 0.161743
## demo_raceOther            0.0207119  0.2431672   0.085 0.932594
## demo_genderWomen          -0.3087733  0.1070130  -2.885 0.006566 **
## cc_smokeFormer            0.0729305  0.1250743   0.583 0.563463
## cc_smokeCurrent           0.2004411  0.1685633   1.189 0.242179
## cc_bmi25 to <30           -0.0215815  0.1116762  -0.193 0.847849
## cc_bmi30 to <35           0.2261152  0.1130356   2.000 0.053045 .
## cc_bmi35+                 0.7023231  0.1345164   5.221 7.63e-06 ***
## cc_diabetesYes            0.3421597  0.1262108   2.711 0.010214 *
## cc_ckdYes                 0.6520772  0.1190915   5.475 3.49e-06 ***
## cc_cvd_anyYes            -0.0273710  0.1208875  -0.226 0.822158
## chol_trig                 0.0017308  0.0003819   4.533 6.21e-05 ***
## chol_ldl                  0.0052070  0.0012622   4.125 0.000209 ***
## chol_hdl                  0.0061692  0.0032180   1.917 0.063190 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasibinomial family taken to be 0.9486429)
##
## Number of Fisher Scoring iterations: 5
```

*#need to look into subset of high cholesterol and hypertension?*

## Modeling medication use with bp\_uncontrolled\_140\_90 outcome

Model is limited to survey design subset of hypertension patients from 2013-2020. We selected medication variables concerning self-reported blood pressure medication use (yes/no), the number of antihypertensive medication pills taken per day, the use of vasodilators (yes/no), alpha and beta blockers (yes/no), the use of angiotensin receptor blockers, and the use of statins or other cholesterol medication for cholesterol control (yes/no). These variables were selected based on their prior association with blood pressure levels in existing clinical trials and research studies.

```
meds_binary <- svyglm(bp_uncontrolled_140_90 ~ bp_med_use + bp_med_n_pills + bp_med_vasod + bp_med_alpha +
summary(meds_binary)
```

```
##
## Call:
## svyglm(formula = bp_uncontrolled_140_90 ~ bp_med_use + bp_med_n_pills +
##       bp_med_vasod + bp_med_alpha + bp_med_beta + bp_med_angioten +
##       chol_med_statin + chol_med_use, design = htsub2, family = quasibinomial())
##
## Survey design:
## subset(htsub, svy_year == "2013-2014" | svy_year == "2015-2016" |
##       svy_year == "2017-2020")
##
## Coefficients:
```

```
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -2.18989    0.06676  -32.800 < 2e-16 ***
## bp_med_useYes       0.47223    0.18528   2.549  0.01437 *
## bp_med_n_pillsOne   0.54676    0.24644   2.219  0.03172 *
## bp_med_n_pillsTwo   0.68321    0.25493   2.680  0.01032 *
## bp_med_n_pillsThree 0.66705    0.30127   2.214  0.03205 *
## bp_med_n_pillsFour or more 0.67142    0.31766   2.114  0.04025 *
## bp_med_vasodYes     0.39992    0.38925   1.027  0.30984
## bp_med_alphaYes     0.08106    0.33994   0.238  0.81263
## bp_med_betaYes      0.18103    0.10856   1.667  0.10252
## bp_med_angiotenYes   0.40086    0.14562   2.753  0.00855 **
## chol_med_statinYes   0.14702    0.39001   0.377  0.70802
## chol_med_useYes      0.01408    0.36954   0.038  0.96978
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasibinomial family taken to be 1.011587)
##
## Number of Fisher Scoring iterations: 4
```

Use of blood pressure medications, number of antihypertensive pills (1-4+), and angiotensin receptor blockers, were significantly associated with having uncontrolled blood pressure. The use of vasodilators, statins or other cholesterol medication, alpha and beta blockers, had non-statistically significant associations with having uncontrolled blood pressure.

### Modeling medication use with bp\_sys\_mean outcome

We now model the continuous outcome of the average systolic blood pressure with the same medication covariates among those with hypertension from 2013-2020.

```
meds_continuous <- svyglm(bp_sys_mean ~ bp_med_use + bp_med_n_pills + bp_med_vasod + bp_med_alpha + bp_med_beta + bp_med_angioten + chol_med_statin + chol_med_use, design = htsub2, family = gaussian())
summary(meds_continuous)
```

```
##
## Call:
## svyglm(formula = bp_sys_mean ~ bp_med_use + bp_med_n_pills +
##       bp_med_vasod + bp_med_alpha + bp_med_beta + bp_med_angioten +
##       chol_med_statin + chol_med_use, design = htsub2, family = gaussian())
##
## Survey design:
## subset(htsub, svy_year == "2013-2014" | svy_year == "2015-2016" |
##       svy_year == "2017-2020")
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      118.6846    0.2625  452.195 < 2e-16 ***
## bp_med_useYes       5.9810    1.3092   4.568 3.95e-05 ***
## bp_med_n_pillsOne   3.0950    1.2975   2.385  0.0214 *
## bp_med_n_pillsTwo   3.0745    1.6920   1.817  0.0760 .
## bp_med_n_pillsThree 3.7637    1.9311   1.949  0.0577 .
## bp_med_n_pillsFour or more 3.1318    3.7009   0.846  0.4020
## bp_med_vasodYes     7.1119    4.7163   1.508  0.1387
```

```

## bp_med_alphaYes          1.1634      3.6998    0.314    0.7547
## bp_med_betaYes           1.7999      1.0073    1.787    0.0808 .
## bp_med_angiotenYes       4.0761      1.3338    3.056    0.0038 **
## chol_med_statinYes       -0.2361      2.6887   -0.088    0.9304
## chol_med_useYes          2.6995      2.5886    1.043    0.3027
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 261.5449)
##
## Number of Fisher Scoring iterations: 2

```

The use of blood pressure medications, number of antihypertensive pills (1), and angiotensin receptor blockers were significantly associated with mean systolic blood pressure levels in those with hypertension from 2013-2020. The use of (2-4+) antihypertensive pills per day, vasodilators, alpha and beta blockers, and statins and other cholesterol medications, had non-statistically significant associations with mean systolic blood pressure levels in those with hypertension from 2013-2020.