Analysis

SW

7/28/2020

```r
library(tidyverse)
library(Hmisc)
library(nhanesA)
library(broom)
library(rstanarm)
library(brms)
library(nnet)
library(effects)
library(survey)
options(mc.cores = 4)
```

Functions

```r
th <- theme_classic() +
    theme(legend.title=element_text(size=18),
        legend.text=element_text(size=15),
        axis.text=element_text(size=15),
        axis.text.x=element_text(angle=90,vjust=.5),
        axis.title=element_text(size=25))

plot_vars <- function(usertype,outcome,vars=NULL,data){
  if (!is.null(vars)) vars <- sort(vars)

  data$users <- data[,usertype] %>% unlist()
  data$oc <- data[,outcome] %>% unlist()

  d <- data %>%
    filter(!is.na(oc),
           !is.na(bmicat),
           !is.na(users)) %>%
    distinct()

  if (is.null(vars)){
    d <- d %>%
      group_by(users,oc) %>%
      dplyr::summarize(f=sum(wtintadj)) %>%
      mutate(f=f/sum(f)) %>%
      ggplot(aes(x=users,y=f,fill=oc)) +
      geom_col(color='black') +
      coord_flip()
```

```r
}else if (identical(vars,'ins')){
  d <- d %>%
    group_by(users,oc,ins) %>%
    dplyr::summarize(f=sum(wtintadj)) %>%
    group_by(users,ins) %>%
    mutate(f=f/sum(f)) %>%
    ggplot(aes(x=users,y=f,fill=oc)) +
    geom_col(color='black') +
    coord_flip() +
    facet_wrap(~ins)
}else if (identical(vars,c('ins','sex'))){
  d <- d %>%
    group_by(users,oc,ins,sex) %>%
    dplyr::summarize(f=sum(wtintadj)) %>%
    group_by(users,ins,sex) %>%
    mutate(f=f/sum(f)) %>%
    ggplot(aes(x=users,y=f,fill=oc)) +
    geom_col(color='black') +
    coord_flip() +
    facet_wrap(sex~ins)
}else if (identical(vars,c('ins','race'))){
  d <- d %>%
    group_by(users,oc,ins,race) %>%
    dplyr::summarize(f=sum(wtintadj)) %>%
    group_by(users,ins,race) %>%
    mutate(f=f/sum(f)) %>%
    ggplot(aes(x=users,y=f,fill=oc)) +
    geom_col(color='black') +
    coord_flip() +
    facet_grid(race~ins)
}else if (identical(vars,c('bmicat','ins'))){
  d <- d %>%
    group_by(users,oc,ins,bmicat) %>%
    dplyr::summarize(f=sum(wtintadj)) %>%
    group_by(users,ins,bmicat) %>%
    mutate(f=f/sum(f)) %>%
    ggplot(aes(x=users,y=f,fill=oc)) +
    geom_col(color='black') +
    coord_flip() +
    facet_grid(bmicat~ins)
}else if (identical(vars,c('ins','race','sex'))){
  d <- d %>%
    group_by(users,oc,ins,race,sex) %>%
    dplyr::summarize(f=sum(wtintadj)) %>%
    group_by(users,ins,race,sex) %>%
    mutate(f=f/sum(f)) %>%
    ggplot(aes(x=users,y=f,fill=oc)) +
    geom_col(color='black') +
    coord_flip() +
    facet_grid(race ~ sex + ins)
}else if (identical(vars,c('bmicat','ins','sex'))){
  d <- d %>%
    group_by(users,oc,ins,sex,bmicat) %>%
```

```r
    dplyr::summarize(f=sum(wtintadj)) %>%
    group_by(users,ins,bmicat,sex) %>%
    mutate(f=f/sum(f)) %>%
    ggplot(aes(x=users,y=f,fill=oc)) +
    geom_col(color='black') +
    coord_flip() +
    facet_grid(bmicat ~ sex + ins)
  }
  d + labs(fill=outcome)
}

get_labs <- function(x, y) {
  labsx <- label(x)
  labsy <- label(y)
  overlap <- c(labsx[labsx != ''], labsy[labsy != ''])
  overlap <- overlap[!duplicated(overlap)]
  if (length(overlap) == 0)
    return(labsx)
  else
    return(overlap)
}

get_data <- function(query,
                     type = NULL,
                     translate = TRUE) {
  out <- NULL
  suffix <- c('', paste0('_', LETTERS[2:10]))
  yrs <- seq(1999, 1999 + 2 * length(suffix), 2)
  for (s in seq_along(suffix)) {
    qname <- paste0(query, suffix[s])
    d  <- suppressMessages(nhanes(qname))
    if (is.null(d))
      next

    d$year <- yrs[s]

    if (is.null(out)) {
      out <-
        as_tibble(d) %>% mutate_if(is.factor, as.character.factor) %>%
        mutate_if(is.integer, as.character)
      labs <- get_labs(out, d)
    } else{
      labs <- get_labs(out, d)
      out <-
        out %>% bind_rows(
          d %>% mutate_if(is.factor, as.character.factor) %>%
            mutate_if(is.integer, as.character)
        )
    }
    update <- rep('', ncol(out))
    names(update) <- colnames(out)
    update[names(labs)] <- labs
    label(out, self = FALSE) <- update
```

```r
  }

  if (translate) {
    if (!is.null(type)) {
      d_vars <- NULL
      for (s in seq_along(suffix)) {
        qname <- paste0(query, suffix[s])
        d_vars_tmp <-
          try(nhanesTableVars(type, qname, namesonly = TRUE), silent = TRUE)
        if (class(d_vars_tmp) == 'try-error')
          next
        update <- setdiff(d_vars_tmp, d_vars)
        if (length(update) > 0) {
          d_vars <- c(update, d_vars)
          out <-
            suppressWarnings(nhanesTranslate(qname, d_vars, data = out))
        }
      }
    }
  }

  return(out)
}


opioids <- c(
  'CODEINE',
  'OXYCODONE',
  'TRAMADOL',
  'MORPHINE',
  'HYDROMORPHONE',
  'FENTANYL',
  'CARFENTANIL',
  'PROPOXYPHENE',
  'HYDROCODONE',
  'HEROIN',
  'PENTAZOCINE',
  'DIHYDROCODEINE',
  'MEPERIDINE',
  'TAPENTADOL',
  'OPIUM'
) # 'BUPRENORPHINE','NALOXONE'
find_drugs <- function(rxs, lookup) {
  lookup <- paste0('\\<', opioids, '\\>')
  hits <- vector(length = length(rxs))
  for (i in seq_along(rxs)) {
    r <- rxs[i]
    if (is.na(r) | r == '')
      next
    for (l in lookup) {
      if (grepl(l, r)) {
        hits[i] <- TRUE
        next
```

```
      }
    }
  }
  return(as.integer(hits))
}
```

Load and Prepare Data

```
dat <-
  readRDS('D:/Downloads/substance_use_analysis/tifdata_paper.rds')

subtract <- function(x, y)
  x - y
cancerdx <- dat$cancer %>%
  select(SEQN, year, starts_with('MCQ240')) %>%
  left_join(dat$demo %>% select(SEQN, year, age = RIDAGEYR),
            by = c('SEQN', 'year')) %>%
  mutate_if(is.factor, as.character.factor) %>%
  mutate_at(vars(starts_with('MCQ240')),
            ~ ifelse(. == "Don't know" | . == '99999' | . == '77777', NA, .)) %>%
  mutate_at(vars(starts_with('MCQ240')),  ~ ifelse(. == '80 years or older', '9999', .)) %>%
  mutate_at(vars(starts_with('MCQ240')), as.integer) %>%
  mutate(age = as.integer(age)) %>%
  mutate_at(vars(starts_with('MCQ240')),  ~ subtract(age, .)) %>%
  mutate_at(vars(starts_with('MCQ240')),  ~ ifelse(. < 1, 1, 0)) %>%
  mutate(cancer1yr = rowSums(select(., starts_with('MCQ240')), na.rm = TRUE)) %>%
  filter(cancer1yr > 0) %>%
  select(SEQN, year, cancer1yr)

opioids <- dat$druginfo %>%
  filter(
    RXDDCI1A == '57' & RXDDCI1B == '58' & RXDDCI1C %in% c('60', '191') |
      RXDDCI2A == '57' &
      RXDDCI2B == '58' & RXDDCI2C %in% c('60', '191') |
      RXDDCI3A == '57' &
      RXDDCI3B == '58' & RXDDCI3C %in% c('60', '191')
  ) %>%
  select(id = RXDDRGID, drug = RXDDRUG)
# filter(!grepl('NALOXONE|BUPRENORPHINE',drug))


drugusers <- dat$otherdruguse %>%
  select(
    SEQN,
    year,
    starts_with('DUQ270'),
    starts_with('DUQ350'),
    starts_with('DUQ310'),
    starts_with('DUQ400'),
    DUQ240,
    DUQ250,
```

```r
    DUQ290,
    DUQ330,
    DUQ370
) %>%
filter(DUQ240 %in% c('Yes', 'No')) %>%
# filter_at(vars(DUQ250,DUQ290,DUQ330,DUQ370),all_vars(. %in% c('Yes','No'))) %>%
mutate_at(vars(ends_with('Q')), as.numeric) %>%
mutate_at(vars(ends_with('Q')),  ~ ifelse(. == 7777 |
                                            . == 9999, NA, .)) %>%
# filter_at(vars(ends_with('Q')),any_vars(!is.na(.))) %>%
mutate(
  days_cocaine = case_when(
    DUQ270U == 'Days' ~ DUQ270Q,
    DUQ270U == 'Weeks' ~ DUQ270Q * 7,
    DUQ270U == 'Months' ~ DUQ270Q * 30,
    DUQ270U == 'Year' ~ DUQ270Q * 365
  ),
  days_heroin = case_when(
    DUQ310U == 'Days' ~ DUQ310Q,
    DUQ310U == 'Weeks' ~ DUQ310Q * 7,
    DUQ310U == 'Months' ~ DUQ310Q * 30,
    DUQ310U == 'Year' ~ DUQ310Q * 365
  ),
  days_meth = case_when(
    DUQ350U == 'Days' ~ DUQ350Q,
    DUQ350U == 'Weeks' ~ DUQ350Q * 7,
    DUQ350U == 'Months' ~ DUQ350Q * 30,
    DUQ350U == 'Year' ~ DUQ350Q * 365
  ),
  days_iv = case_when(
    DUQ400U == 'Days' ~ DUQ400Q,
    DUQ400U == 'Weeks' ~ DUQ400Q * 7,
    DUQ400U == 'Months' ~ DUQ400Q * 30,
    DUQ400U == 'Year' ~ DUQ400Q * 365
  )
) %>%
rowwise() %>%
mutate(drugdur = min(days_cocaine, days_heroin, days_meth, days_iv, na.rm =
                        TRUE)) %>%
ungroup() %>%
mutate(
  drugdurcat = case_when(
    drugdur <= 365 ~ '<1y',
    drugdur > 365 & drugdur < Inf ~ '>1y',
    DUQ240 == 'No' |
      (DUQ250 == 'No' &
         DUQ290 == 'No' & DUQ330 == 'No' & DUQ370 == 'No') ~ 'none'
  )
) %>%
# filter(!is.na(drugdurcat)) %>%
mutate(druguse = ifelse(drugdurcat == 'none', 'No', 'Yes')) %>%
select(SEQN, year, druguse, drugdur, drugdurcat, starts_with('days'))
```

```r
opioidusers <- dat$druguse %>%
  select(
    SEQN,
    year,
    rxcount = RXD295,
    rxuse = RXDUSE,
    rx = RXDDRUG,
    rxid = RXDDRGID,
    rxdur = RXDDAYS
  ) %>%
  # filter(opioiduse %in% c('Yes','No')) %>%
  mutate(
    opioiduse = case_when((rxid %in% opioids$id) ~ 'Yes',!(rxid %in% opioids$id) |
                             rxuse == 'No' ~ 'No'
    ),
    opioiddur = ifelse(opioiduse == 'Yes', as.integer(rxdur), NA),
    opioiddurcat = case_when(
      opioiduse == 'No' ~ 'none',
      opioiddur < 90 ~ 'short',
      opioiddur >= 90 &
        opioiddur <= 25550 ~ 'long'
    )
  ) %>%
  mutate(
    rxcount = as.integer(rxcount),
    rxcount_listed = str_count(as.character.factor(rx), ';')
  )


opioidusers <- dat$druguse %>%
  select(
    SEQN,
    year,
    rxcount = RXD295,
    rxuse = RXDUSE,
    rx = RXDDRUG,
    rxid = RXDDRGID,
    rxdur = RXDDAYS
  ) %>%
  # filter(opioiduse %in% c('Yes','No')) %>%
  mutate(
    opioiduse = case_when((rxid %in% opioids$id) ~ 'Yes',!(rxid %in% opioids$id) |
                             rxuse == 'No' ~ 'No'
    ),
    opioiddur = ifelse(opioiduse == 'Yes', as.integer(rxdur), NA),
    rxcount = as.integer(rxcount),
    rxcount_listed = str_count(as.character.factor(rx), ';'),
    rxcount_listed = ifelse(
      nchar(as.character.factor(rx)) > 0,
      rxcount_listed + 1,
      rxcount_listed
    )
  ) %>%
```

```r
  group_by(SEQN, year) %>%
  mutate(
    rxcount_listed = sum(rxcount_listed),
    opioiduse = ifelse(any(opioiduse == 'Yes'), 'Yes', opioiduse),
    opioiddur = ifelse(opioiddur == 99999 |
                         opioiddur == 77777, NA, opioiddur),
    opioiddur = max(opioiddur, na.rm = TRUE),
    opioiddur = ifelse(is.infinite(opioiddur), NA, opioiddur)
  ) %>%
  ungroup() %>%
  nest(rx = c(rx, rxdur, rxid)) %>%
  mutate(
    opioiddurcat = case_when(
      opioiduse == 'No' ~ 'none',
      opioiddur < 90 ~ 'short',
      opioiddur >= 90 &
        opioiddur <= 25550 ~ 'long'
    )
  )


users <- full_join(drugusers, opioidusers, by = c('SEQN', 'year')) %>%
  filter((druguse %in% c('Yes', 'No')) |
           (opioiduse %in% c('Yes', 'No')),!is.na(druguse) |
           !is.na(opioiduse)) %>%
  left_join(
    dat$smoke %>% # might be wrong
      select(SEQN, year, smokecurrent = SMQ040, smokelife = SMQ020) %>%
      mutate(
        smoke = case_when(
          smokecurrent %in% c('Every day,', 'Some days, or') ~ 'Current',
          smokelife == 'Yes' ~ 'Former',
          smokecurrent == 'Not at all?' |
            smokelife == 'No' ~ 'Never'
        )
      ) %>%
      select(SEQN, year, smoke),
    by = c('SEQN', 'year')
  ) %>%
  left_join(
    dat$ins %>%
      select(SEQN, year, ins1 = HID010, ins2 = HIQ011) %>%
      mutate(
        ins = case_when(ins1 == 'Yes' | ins2 == 'Yes' ~ 'Yes',
                        ins1 == 'No' |
                          ins2 == 'No' ~ 'No')
      ) %>%
      select(SEQN, year, ins),
    by = c('SEQN', 'year')
  ) %>%
  left_join(
    dat$demo %>% select(
      SEQN,
```

```
    year,
    age = RIDAGEYR,
    sex = RIAGENDR,
    race = RIDRETH1,
    edu = DMDEDUC2
  ) %>%
    mutate(
      age = as.integer(age),
      edu = case_when(
        edu %in% c('Less Than 9th Grade',
                   '9-11th Grade (Includes 12th grad') ~ 'Less than high school' ,
        edu == 'High School Grad/GED or Equivale' ~ 'High school or equivalent',
        edu == 'Some College or AA degree' ~ 'Some college',
        edu == 'College Graduate or above' ~ 'College or higher'
      ),
      race = case_when(
        race %in% c('Mexican American', 'Other Hispanic') ~ 'Hispanic',
        race == 'Other Race - Including Multi-Rac' ~ 'Non-Hispanic Other',
        race == 'Non-Hispanic White' ~ 'Non-Hispanic White',
        race == 'Non-Hispanic Black' ~ 'Non-Hispanic Black'
      )
    ),
  by = c('SEQN', 'year')
) %>%
left_join(
  dat$demo_raw %>% select(
    SEQN,
    cycle = SDDSRVYR,
    year,
    wtint2yr = WTINT2YR,
    wtmec2yr = WTMEC2YR,
    cluster = SDMVPSU,
    strata = SDMVSTRA
  ),
  by = c('year', 'SEQN')
) %>%
filter(age >= 18 & age <= 100) %>%
left_join(dat$body %>% select(SEQN, year, bmi = BMXBMI), by = c('SEQN', 'year')) %>%
mutate(
  bmicat = case_when(
    bmi < 20 ~ 'underweight',
    bmi >= 20 & bmi < 25 ~ 'normal',
    bmi >= 25 & bmi < 30 ~ 'overweight',
    bmi >= 30 & bmi < 35 ~ 'obese I',
    bmi >= 35 & bmi < 40 ~ 'obese II',
    bmi >= 40 & bmi <= 80 ~ 'obese III'
  ),
  agecat = case_when(
    age >= 18 & age < 25 ~ '18-24',
    age >= 25 & age < 35 ~ '25-34',
    age >= 35 & age < 45 ~ '35-44',
    age >= 45 & age < 55 ~ '45-54',
    age >= 55 & age < 65 ~ '55-64',
    age >= 65 & age < 75 ~ '65-74',
```

```r
    age >= 75 & age < 100 ~ '75-100'
  )
) %>%
left_join(dat$preg %>% select(SEQN, year, preg = RHD143),
          by = c('SEQN', 'year')) %>%
mutate(preg = case_when(preg == 'Yes' ~ 'Yes', TRUE ~ 'No')) %>%
left_join(cancerdx, by = c('SEQN', 'year')) %>%
filter(preg != 'Yes' | is.na(preg)) %>%
filter(is.na(cancer1yr)) %>%
filter(!is.na(smoke)) %>%
filter(!is.na(edu)) %>%
filter(ins %in% c('Yes', 'No')) %>%
mutate(
  sex = relevel(as.factor(sex), ref = 'Male'),
  opioiddurcat = relevel(as.factor(opioiddurcat), ref = 'none'),
  drugdurcat = relevel(as.factor(drugdurcat), ref = 'none'),
  bmicat = relevel(as.factor(bmicat), ref = 'normal'),
  agecat = relevel(as.factor(agecat), ref = '35-44'),
  race = relevel(as.factor(race), ref = 'Non-Hispanic White'),
  edu = relevel(as.factor(edu), ref = 'High school or equivalent'),
  ins = relevel(as.factor(ins), ref = 'No'),
  smoke = relevel(as.factor(smoke), ref = 'Never'),
  yearcat = relevel(as.factor(year), ref = '2003')
) %>%
mutate(
  strata = dense_rank(strata),
  cluster = dense_rank(cluster),
  opioiddurrank = case_when(
    opioiddurcat == 'none' ~ 1,
    opioiddurcat == 'short' ~ 2,
    opioiddurcat == 'long' ~ 3
  )
) %>%
select(
  SEQN,
  year,
  yearcat,
  bmi,
  bmicat,
  age,
  agecat,
  sex,
  race,
  edu,
  ins,
  smoke,
  strata,
  cluster,
  rx,
  rxcount,
  rxcount_listed,
  drugdur,
  drugdurcat,
```

```r
    druguse,
    starts_with('days'),
    opioiddur,
    opioiddurcat,
    opioiddurrank,
    opioiduse,
    wtint2yr,
    wtmec2yr
  ) %>%
  mutate(wtintadj = wtint2yr / length(unique(year)),
         wtmecadj = wtmec2yr / length(unique(year)))


dat_everything <- readRDS('/Research/Tiffany/dat_everything.rds')
dat_everything <- dat_everything %>%
  select(
    union(
      dat_everything %>%
        filter(SEQN %in% (
          users %>% filter(opioiduse == 'Yes') %>% select(SEQN) %>% distinct() %>% unlist()
        )) %>%
        select_if( ~ sum(!is.na(.)) > 100) %>% # consider changing
        colnames(),
      dat_everything %>%
        filter(SEQN %in% (
          users %>% filter(druguse == 'Yes') %>% select(SEQN) %>% distinct() %>% unlist()
        )) %>%
        select_if( ~ sum(!is.na(.)) > 100) %>% # consider changing
        colnames()
    )
  ) %>%
  select(-BMXBMI) %>%
  mutate_if(is.factor, as.character.factor) %>%
  mutate_all( ~ ifelse(. %in% c('Refused', "Don't know"), NA, .))

users_everything <- users %>%
  left_join(dat_everything, by = c('SEQN', 'year')) %>%
  rename(
    normal_otoscopy_L = AUXOTSPL,
    normal_otoscopy_R = AUXROTSP,
    middle_ear_pres_R = AUXTMEPR,
    middle_ear_pres_L = AUXTMEPL,
    tympanogram_qual_R = AUAREQC,
    tympanogram_qual_L = AUALEQC,
    hear_gen = AUQ130,
    hear_last_test = AUQ140,
    hear_aid = AUQ150,
    hear_recent_sx = AUQ190,
    pulse = BPXPLS,
    bp_sys_1 = BPXSY1,
    bp_sys_2 = BPXSY2,
    bp_sys_3 = BPXSY3,
    bp_sys_4 = BPXSY4,
```

```
bp_di_1 = BPXDI1,
bp_di_2 = BPXDI2,
bp_di_3 = BPXDI3,
bp_di_4 = BPXDI4,
bp_high = BPQ020,
bp_med = BPQ040A,
chol_high = BPQ080,
chol_med = BPQ090D,
wt = BMXWT,
ht = BMXHT,
overweight1 = MCQ080,
overweight2 = MCQ160J,
feno = ENXMEAN,
fev1_bl = SPXNFEV1,
fet_bl = SPXNFET,
fvc_bl = SPXNFVC,
vis_R = VIDRVA,
vis_L = VIDLVA,
trouble_seeing = MCQ140,
vis_acuity_R = VIDROVA,
vis_acuity_L = VIDLOVA,
gen_health1 = HSD010,
days_in_month_poor_health = HSQ470,
days_in_month_poor_mental = HSQ480,
days_in_month_forced_inactive = HSQ490,
cp = CDQ001,
cp_exert = CDQ002,
cp_noexert = CDQ003,
cp_severe30 = CDQ008,
sob_stairs = CDQ010,
dm = DIQ010,
dm_insulin = DIQ050,
dm_med = DIQ070,
dm_eye = DIQ080,
dm_ulcer = DIQ090,
dm_numb_recent_handfeet = DIQ100,
dm_numb_loc = DIQ110,
dm_paintingle_recent_handfeet = DIQ120,
dm_paintingle_loc = DIQ130,
dm_pain_leg = DIQ140,
dm_pain_calf = DIQ150,
run_out_food = FSD032A,
emo_support = SSQ011,
hepa_vax = IMD010,
hepb_vax = IMQ020,
liverdz1 = MCQ160L,
liverdz2 = MCQ170L,
liverdz_age = MCQ180L,
liver_cap_med = LUXCAPM,
liver_cap_iqr = LUXCPIQR,
pneumo_vax = IMQ030,
asthma = MCQ010,
asthma_recent_treatment = MCQ053,
```

```
asthma_recent_attack = MCQ040,
asthma_recent_er = MCQ050,
emphysema = MCQ160G,
emphysema_age = MCQ180G,
chr_bronch1 = MCQ160K,
chr_bronch2 = MCQ170K,
chr_bronch_age = MCQ180K,
copd = MCQ160O,
cough_chronic = RDQ031,
cough_prod = RDQ050,
wheeze = RDQ070,
wheeze_med = RDQ134,
cough_nocturnal_dry = RDQ140,
blood_trans = MCQ092,
arthritis = MCQ160A,
arthritis_age = MCQ180A,
gout = MCQ160N,
gout_age = MCQ180N,
fx_hip = OSQ010A,
fx_wrist = OSQ010B,
fx_spine = OSQ010C,
op = OSQ060,
op_tx = OSQ070,
chf = MCQ160B,
chf_age = MCQ180B,
chd = MCQ160C,
chd_age = MCQ180C,
angina = MCQ160D,
angina_age = MCQ180D,
mi = MCQ160E,
mi_age = MCQ180E,
stroke = MCQ160F,
stroke_age = MCQ180F,
thyroiddz1 = MCQ160M,
thyroiddz2 = MCQ160I,
thyroiddz_current = MCQ170M,
thyroiddz_age = MCQ180M,
goiter = MCQ160H,
cancer = MCQ220,
cancer_type = MCQ230A,
dentist_last_visit = OHQ030,
insecure_mouth_freq = OHQ630,
insecure_mouth_freq_recent = OHQ680,
dentist_could_not_get_recent = OHQ770,
herpes = SXQ260,
warts = SXQ265,
gono = SXQ270,
chlam = SXQ272,
hysterectomy = RHD280,
gen_health2 = HUQ010,
gen_health_comp = HUQ020,
hc_loc = HUQ030,
hc_type = HUQ040,
```

```
    hc_count = HUQ050,
    hc_last = HUQ060,
    hosp_nights = HUQ071,
    hosp_nights_count = HUD080,
    mental_visit_count = HUQ090,
    activity_walkbike = PAD020,
    activity_avg_amount = PAQ180,
    smoke_age_reg = SMD055,
    smoke_age = SMD030,
    smoke_num_cigs_per_day_now = SMD070,
    smoke_num_cigs_per_day_now_length = SMD075,
    smoke_num_cigs_per_day = SMD057,
    smoke_avg_cigs_per_day_recent = SMD650,
    smoke_100cigs = SMQ020
) %>%
mutate(
  gen_health1 = gsub(',.*|\\?.*', '', gen_health1, ignore.case = TRUE),
  gen_health2 = gsub(',.*|\\?.*', '', gen_health2, ignore.case =
                      TRUE),
  dm = relevel(factor(dm, ordered = FALSE), ref = 'No'),
  gen_health1 = factor(gen_health1, ordered = TRUE, levels = rev(
    c('Excellent', 'Very good', 'Good', 'Fair', 'Poor')
  )),
  gen_health2 = factor(gen_health2, ordered = TRUE, levels = rev(
    c('Excellent', 'Very good', 'Good', 'Fair', 'Poor')
  )),
  bmicat = factor(
    bmicat,
    ordered = TRUE,
    levels = c(
      'underweight',
      'normal',
      'overweight',
      'obese I',
      'obese II',
      'obese III'
    )
  ),
  agecat = factor(
    agecat,
    ordered = TRUE,
    levels = c('18-24', '25-34', '35-44', '45-54', '55-64', '65-74', '75-100')
  ),
  edu = factor(
    edu,
    ordered = TRUE,
    levels = c(
      'Less than high school',
      'High school or equivalent',
      'Some college',
      'College or higher'
    )
  ),
```

```r
    yearrank = dense_rank(year),
    bmirank = dense_rank(bmicat),
    edurank = dense_rank(edu),
    agerank = dense_rank(agecat)
) %>%
mutate_at(vars(starts_with('bp_sys')), as.integer) %>%
mutate_at(vars(starts_with('bp_di')), as.integer) %>%
rowwise() %>%
mutate(map_mean = mean(c((bp_sys_1 + 2 * bp_di_1) / 3,
                         (bp_sys_2 + 2 * bp_di_2) / 3,
                         (bp_sys_3 + 2 * bp_di_3) / 3
), na.rm = TRUE),
map_min = min(c((bp_sys_1 + 2 * bp_di_1) / 3,
               (bp_sys_2 + 2 * bp_di_2) / 3,
               (bp_sys_3 + 2 * bp_di_3) / 3
), na.rm = TRUE)) %>%
ungroup() %>%
mutate_at(vars(starts_with('map')),  ~ ifelse(is.infinite(.) |
                                              is.nan(.), NA, .))
```

```r
users_everything %>%
  filter(!is.na(hc_loc),!is.na(drugdurcat),!is.na(ins),!is.na(race)) %>%
  group_by(race, drugdurcat, ins, hc_loc) %>%
  dplyr::summarize(f = sum(wtintadj)) %>%
  mutate(f = f / sum(f)) %>%
  ggplot(aes(x = drugdurcat, y = f, fill = hc_loc)) +
  geom_col(color = 'black') +
  coord_flip() +
  facet_grid(race ~ ins) +
  th +
  labs(y = '',
       x = 'Time Since Drug Use',
       fill = 'Healthcare Location',
       title = 'Health Insurance')
```
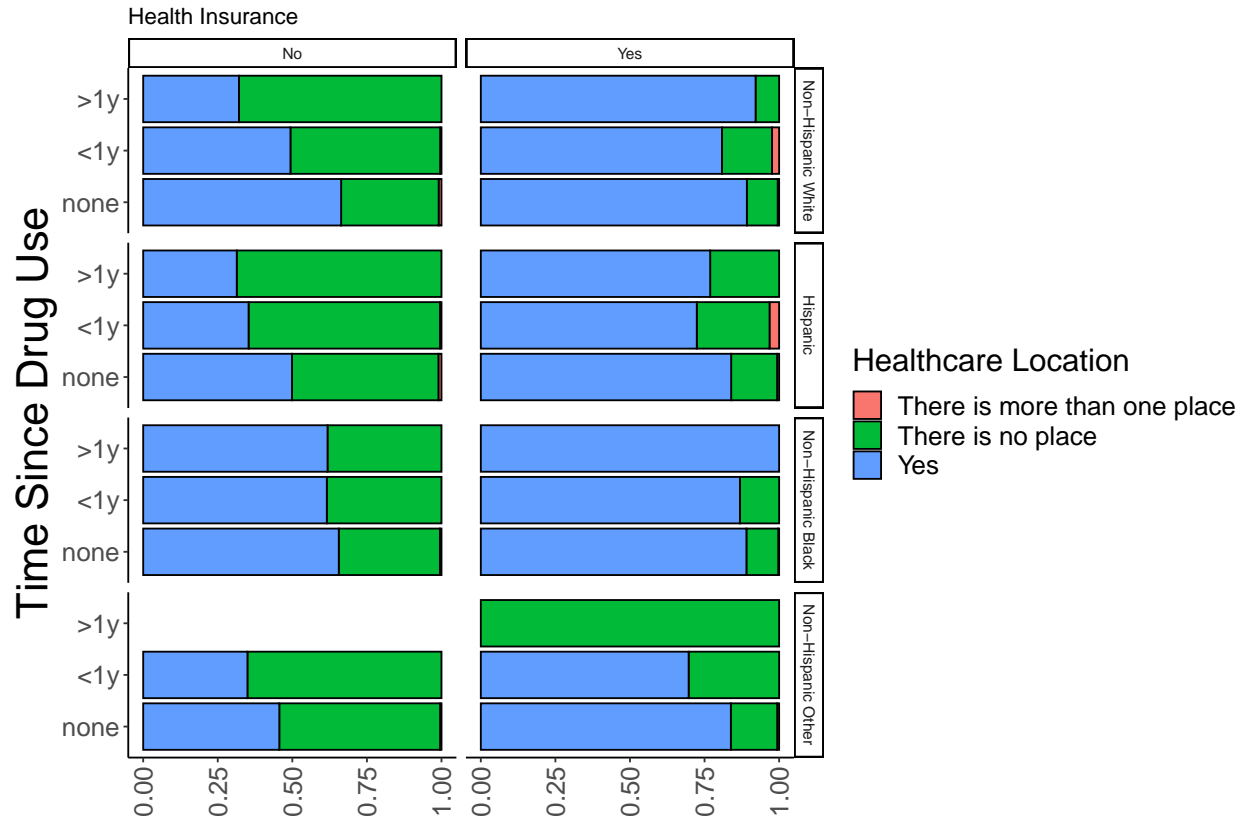
```
users_everything %>%
  filter(!is.na(hc_loc),!is.na(opioiddurcat),!is.na(ins),!is.na(race)) %>%
  group_by(race, opioiddurcat, ins, hc_loc) %>%
  dplyr::summarize(f = sum(wtintadj)) %>%
  mutate(f = f / sum(f)) %>%
  ggplot(aes(x = opioiddurcat, y = f, fill = hc_loc)) +
  geom_col(color = 'black') +
  coord_flip() +
  facet_grid(race ~ ins) +
  th +
  labs(y = '',
       x = 'Duration Using Opioids',
       fill = 'Healthcare Location',
       title = 'Health Insurance')
```
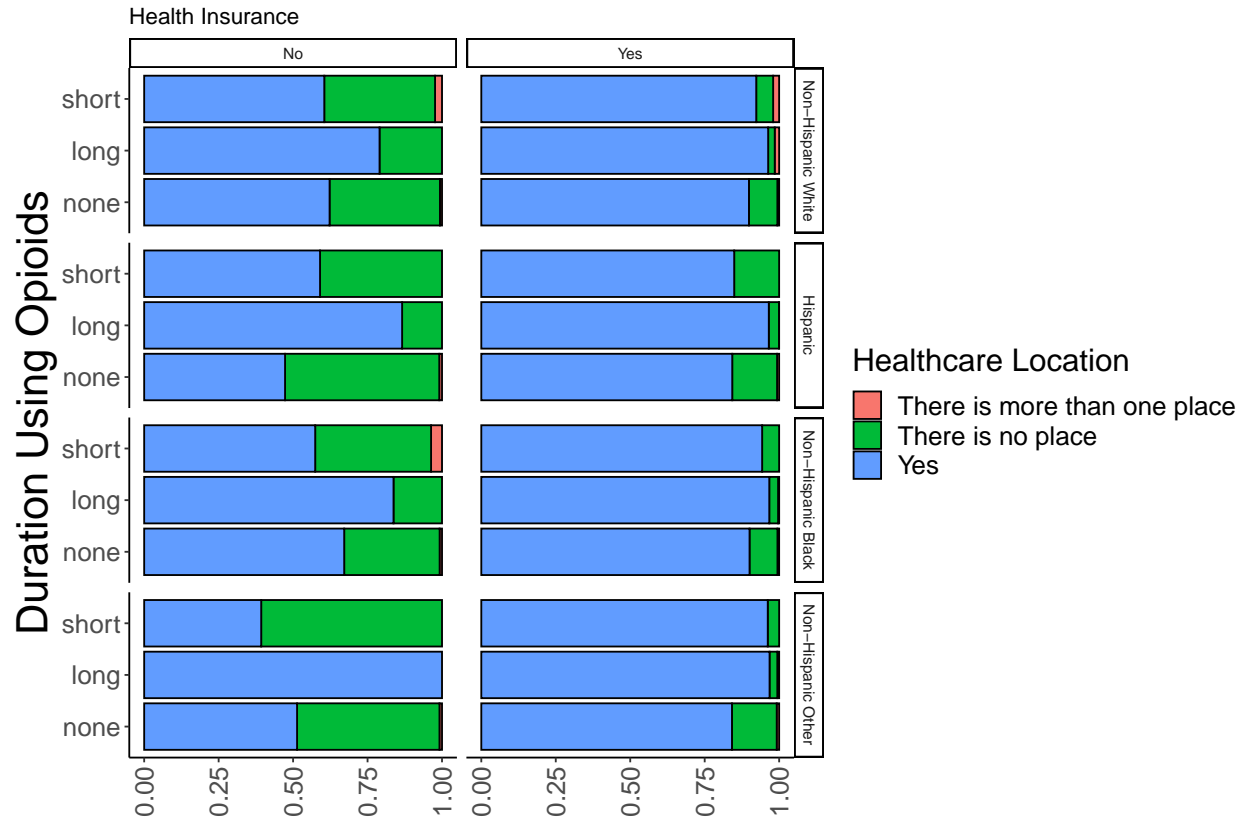
```
users_everything %>%
  filter(!is.na(rxcount_listed),!is.na(opioiddurcat),!is.na(ins),!is.na(agecat)) %>%
  group_by(opioiddurcat, ins, agecat) %>%
  dplyr::summarize(mean = sum(wtintadj * rxcount_listed, na.rm = TRUE) /
                     sum(wtintadj, na.rm = TRUE)) %>%
  ggplot(aes(x = opioiddurcat, y = mean)) +
  geom_col(color = 'black') +
  coord_flip() +
  facet_grid(agecat ~ ins) +
  th +
  labs(y = 'Average Number of Opioids',
       x = 'Duration Using Opioids',
       fill = 'Healthcare Location',
       title = 'Health Insurance')
```
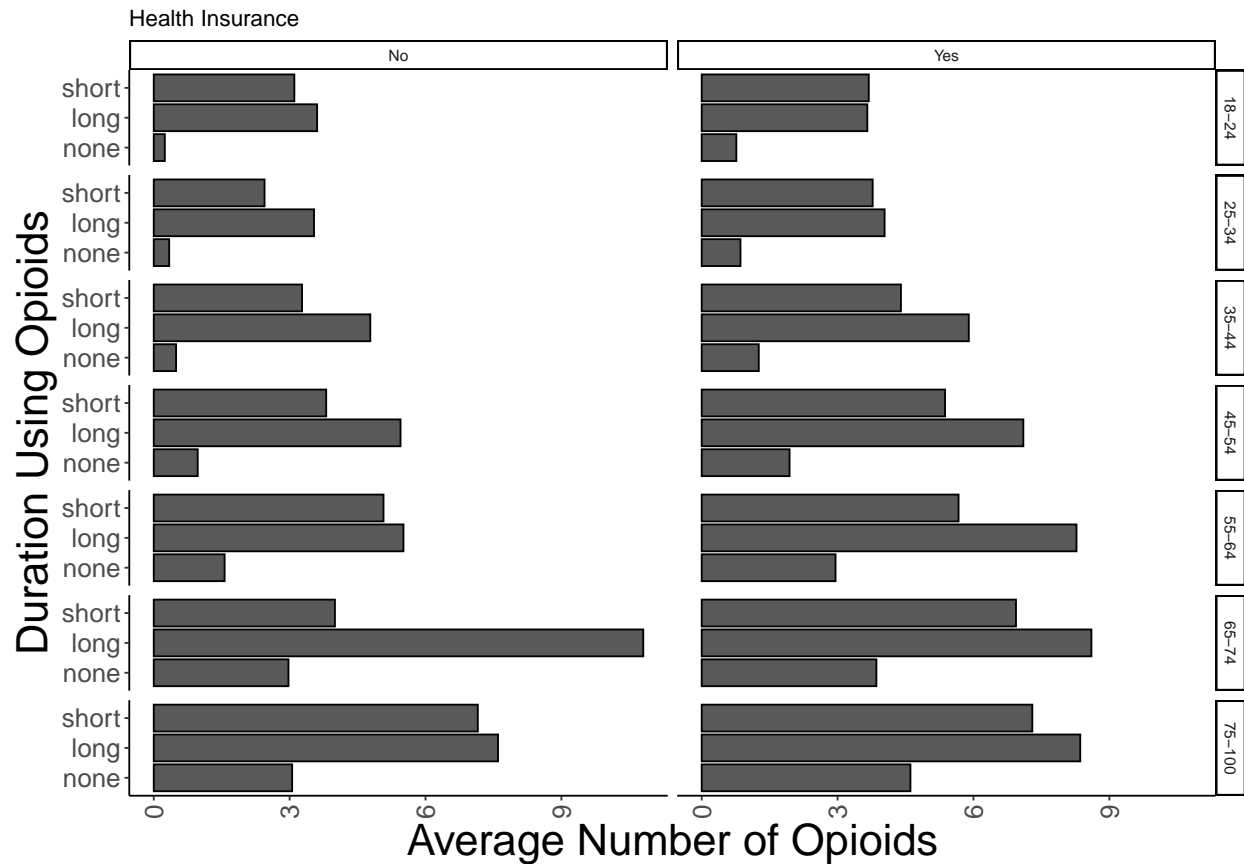
```
users_everything %>%
  filter(!is.na(rxcount_listed),!is.na(drugdurcat),!is.na(ins),!is.na(agecat)) %>%
  group_by(drugdurcat, ins, agecat) %>%
  dplyr::summarize(mean = sum(wtintadj * rxcount_listed, na.rm = TRUE) /
                   sum(wtintadj, na.rm = TRUE)) %>%
  ggplot(aes(x = drugdurcat, y = mean)) +
  geom_col(color = 'black') +
  coord_flip() +
  facet_grid(agecat ~ ins) +
  th +
  labs(y = 'Average Number of Opioids',
       x = 'Time Since Drug Use',
       fill = 'Healthcare Location',
       title = 'Health Insurance')
```
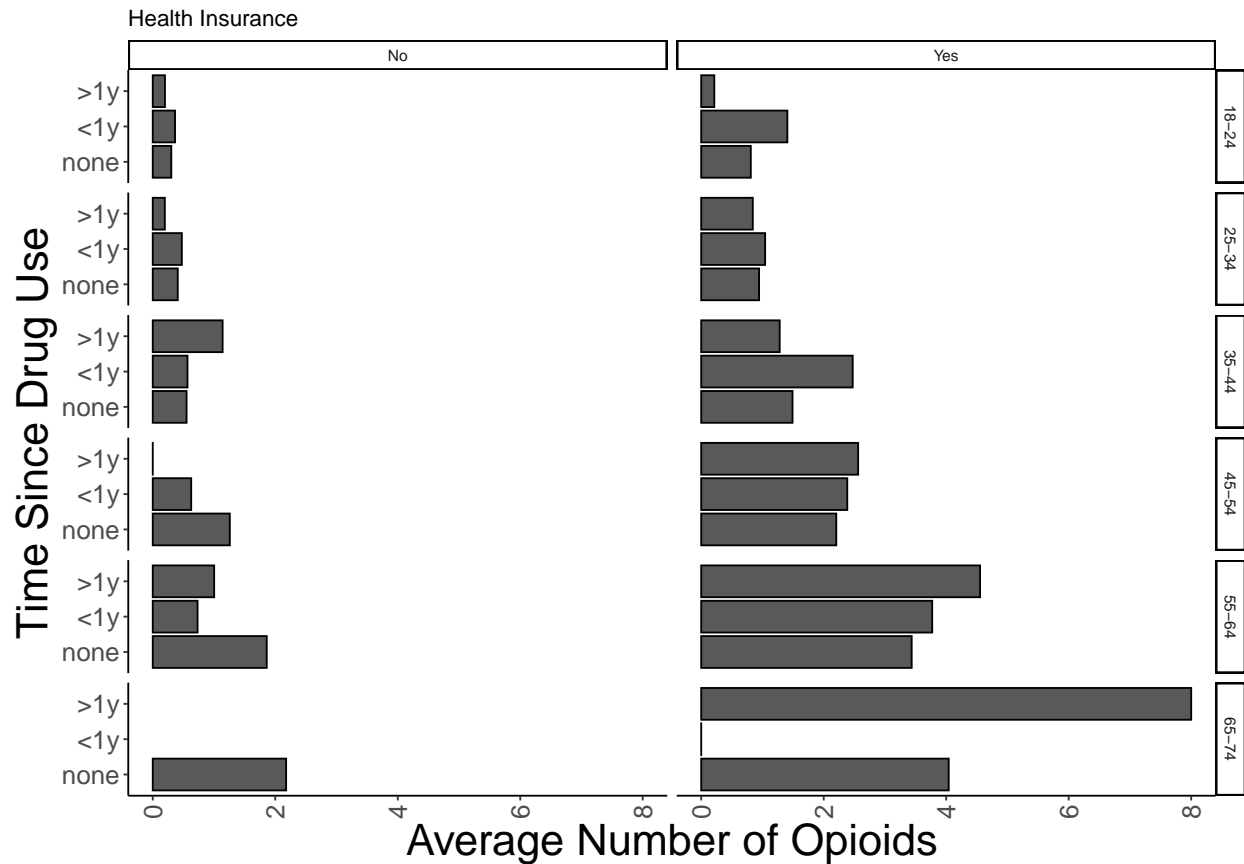
```
users_everything %>%
  filter(!is.na(opioiddurcat),
         !is.na(bp_med)) %>%
  distinct() %>%
  group_by(ins,sex,bp_high,opioiddurcat,bp_med) %>%
  dplyr::summarize(f=sum(wtmecadj)) %>%
  mutate(f=f/sum(f)) %>%
  ggplot(aes(x=opioiddurcat,y=f,fill=bp_med)) +
  geom_col(color='black') +
  coord_flip() +
  facet_grid(~sex+ins) +
  th +
    labs(y = 'Proportion of Subjects',
      x = 'Duration Using Opioids',
      fill = 'On BP Med.',
      title = 'Sex vs Health Insurance')
```
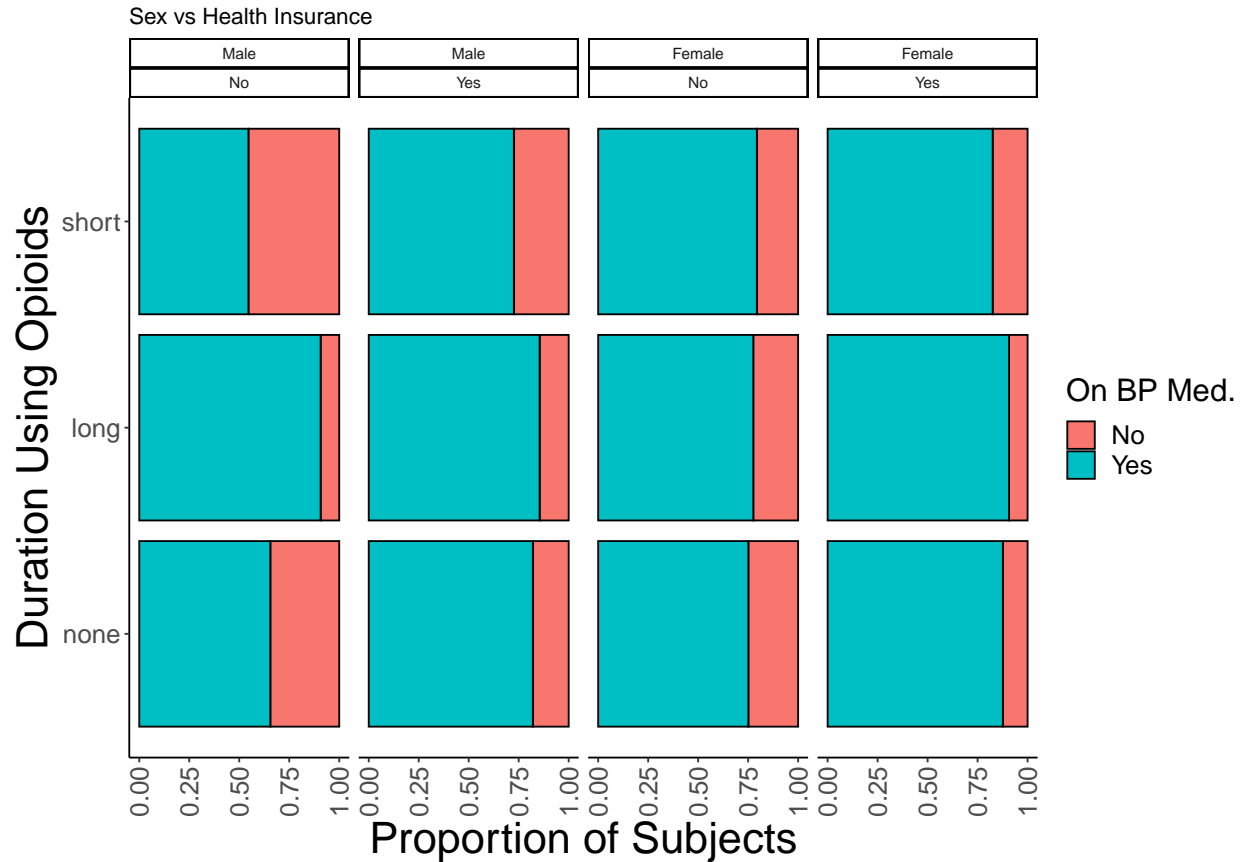
A few sanity checks. First, the impact on a bunch of predictors on chest pain. We're basically controlling for sex, race, age, education, insurance, smoking, year surveyed, and bmi, and then focusing on the impact that the duration of opioid use (short vs long) has on chest pain. I'm using a package that accounts for survey designs as well, with the appropriate weighting. I'd prefer to do this fully Bayesian with post stratification (see the one script), but that's for down the road once we identify some questions of interest and, by extension, interesting analyses designs.

```
svy <- svydesign(
  id =  ~ cluster,
  strata =  ~ strata,
  weights =  ~ wtmecadj,
  data = users_everything %>%
    select(
      SEQN,
      cp,
      sex,
      race,
      agerank,
      edurank,
      ins,
      smoke,
      yearrank,
      bmirank,
      wtmecadj,
      cluster,
      strata,
```

```
      drugdurcat
    ) %>%
    filter(!is.na(cp), !is.na(bmirank), !is.na(drugdurcat)) %>%
    mutate(cp = as.integer(ifelse(cp == 'Yes', 1, 0))) %>%
    distinct(),
  nest = TRUE
)
m <-
  svyglm(
    cp ~ sex + race + agerank + edurank + ins + smoke + yearrank + bmirank + drugdurcat,
    design = svy,
    family = binomial(link = 'logit'),
    rescale = TRUE
  )
summary(m)
```

```
FALSE
FALSE Call:
FALSE svyglm(formula = cp ~ sex + race + agerank + edurank + ins +
FALSE     smoke + yearrank + bmirank + drugdurcat, design = svy, family = binomial(link = "logit"),
FALSE     rescale = TRUE)
FALSE
FALSE Survey design:
FALSE svydesign(id = ~cluster, strata = ~strata, weights = ~wtmecadj,
FALSE     data = users_everything %>% select(SEQN, cp, sex, race, agerank,
FALSE         edurank, ins, smoke, yearrank, bmirank, wtmecadj, cluster,
FALSE         strata, drugdurcat) %>% filter(!is.na(cp), !is.na(bmirank),
FALSE         !is.na(drugdurcat)) %>% mutate(cp = as.integer(ifelse(cp ==
FALSE         "Yes", 1, 0))) %>% distinct(), nest = TRUE)
FALSE
FALSE Coefficients:
FALSE                         Estimate Std. Error t value Pr(>|t|)
FALSE (Intercept)            -2.041799   0.226965  -8.996 2.13e-14 ***
FALSE sexFemale               0.074333   0.054926   1.353  0.17913
FALSE raceHispanic           -0.163271   0.066005  -2.474  0.01513 *
FALSE raceNon-Hispanic Black  0.020245   0.060147   0.337  0.73716
FALSE raceNon-Hispanic Other  0.166974   0.102023   1.637  0.10498
FALSE agerank                 0.072884   0.032275   2.258  0.02620 *
FALSE edurank                -0.083373   0.028615  -2.914  0.00444 **
FALSE insYes                  0.157229   0.069640   2.258  0.02623 *
FALSE smokeCurrent            0.581179   0.071149   8.169 1.25e-12 ***
FALSE smokeFormer             0.271281   0.073944   3.669  0.00040 ***
FALSE yearrank                0.001165   0.013247   0.088  0.93010
FALSE bmirank                 0.148787   0.024605   6.047 2.82e-08 ***
FALSE drugdurcat<1y           0.468560   0.184276   2.543  0.01260 *
FALSE drugdurcat>1y           0.037408   0.547192   0.068  0.94564
FALSE ---
FALSE Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
FALSE
FALSE (Dispersion parameter for binomial family taken to be 0.9997766)
FALSE
FALSE Number of Fisher Scoring iterations: 4
```

Both opioid durations affect the frequency chest pain was reported. The estimate is over twice as large for

long durations (> 90 days), suggesting that the longer you're on opioids, the more likely you are to report chest pain, which is reasonable.

```r
svy <- svydesign(
  id =   ~ cluster,
  strata =   ~ strata,
  weights =   ~ wtmecadj,
  data = users_everything %>%
    select(
      SEQN,
      bp_high,
      sex,
      race,
      agerank,
      edurank,
      ins,
      smoke,
      yearrank,
      bmirank,
      wtmecadj,
      cluster,
      strata,
      opioiddurcat
    ) %>%
    filter(!is.na(bp_high),!is.na(bmirank),!is.na(opioiddurcat)) %>%
    mutate(bp_high = as.integer(ifelse(bp_high == 'Yes', 1, 0))) %>%
    distinct(),
  nest = TRUE
)
m <-
  svyglm(
    bp_high ~ sex + race + agerank + edurank + ins + smoke + yearrank + bmirank + opioiddurcat,
    design = svy,
    family = binomial(link = 'logit'),
    rescale = TRUE
  )
summary(m)
```

```
FALSE
FALSE Call:
FALSE svyglm(formula = bp_high ~ sex + race + agerank + edurank + ins +
FALSE     smoke + yearrank + bmirank + opioiddurcat, design = svy,
FALSE     family = binomial(link = "logit"), rescale = TRUE)
FALSE
FALSE Survey design:
FALSE svydesign(id = ~cluster, strata = ~strata, weights = ~wtmecadj,
FALSE     data = users_everything %>% select(SEQN, bp_high, sex, race,
FALSE         agerank, edurank, ins, smoke, yearrank, bmirank, wtmecadj,
FALSE         cluster, strata, opioiddurcat) %>% filter(!is.na(bp_high),
FALSE         !is.na(bmirank), !is.na(opioiddurcat)) %>% mutate(bp_high = as.integer(ifelse(bp_high ==
FALSE         "Yes", 1, 0))) %>% distinct(), nest = TRUE)
FALSE
FALSE Coefficients:
FALSE                      Estimate Std. Error t value Pr(>|t|)
FALSE (Intercept)          -4.776493   0.084307 -56.656  < 2e-16 ***
```

22

```
FALSE sexFemale                    -0.063372   0.029817  -2.125 0.035311 *
FALSE raceHispanic                 -0.210720   0.043348  -4.861 3.09e-06 ***
FALSE raceNon-Hispanic Black  0.519081   0.034875  14.884  < 2e-16 ***
FALSE raceNon-Hispanic Other  0.175613   0.054056   3.249 0.001451 **
FALSE agerank                       0.574758   0.010139  56.688  < 2e-16 ***
FALSE edurank                      -0.104763   0.014847  -7.056 7.25e-11 ***
FALSE insYes                        0.210399   0.044166   4.764 4.69e-06 ***
FALSE smokeCurrent                  0.225586   0.034855   6.472 1.51e-09 ***
FALSE smokeFormer                   0.118547   0.031858   3.721 0.000286 ***
FALSE yearrank                      0.015727   0.006979   2.253 0.025788 *
FALSE bmirank                       0.451102   0.010736  42.017  < 2e-16 ***
FALSE opioiddurcatlong              0.586182   0.069623   8.419 4.02e-14 ***
FALSE opioiddurcatshort             0.179190   0.122659   1.461 0.146287
FALSE ---
FALSE Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
FALSE
FALSE (Dispersion parameter for binomial family taken to be 0.9861899)
FALSE
FALSE Number of Fisher Scoring iterations: 4
```

Here, only long term opioid use is associated with reporting high blood pressure. The effect is less than what we saw for chest pain, too.

```
svy <- svydesign(
  id =   ~ cluster,
  strata =   ~ strata,
  weights =   ~ wtmecadj,
  data = users_everything %>%
    select(
      SEQN,
      bp_high,
      sex,
      race,
      agerank,
      edurank,
      ins,
      smoke,
      yearrank,
      bmirank,
      wtmecadj,
      cluster,
      strata,
      drugdurcat
    ) %>%
    filter(!is.na(bp_high), !is.na(bmirank), !is.na(drugdurcat)) %>%
    mutate(bp_high = as.integer(ifelse(bp_high == 'Yes', 1, 0))) %>%
    distinct(),
  nest = TRUE
)
m <-
  svyglm(
    bp_high ~ sex + race + agerank + edurank + ins + smoke + yearrank + bmirank + drugdurcat,
    design = svy,
    family = binomial(link = 'logit'),
    rescale = TRUE
```

```
)
summary(m)
```

```
FALSE
FALSE Call:
FALSE svyglm(formula = bp_high ~ sex + race + agerank + edurank + ins +
FALSE     smoke + yearrank + bmirank + drugdurcat, design = svy, family = binomial(link = "logit"),
FALSE     rescale = TRUE)
FALSE
FALSE Survey design:
FALSE svydesign(id = ~cluster, strata = ~strata, weights = ~wtmecadj,
FALSE     data = users_everything %>% select(SEQN, bp_high, sex, race,
FALSE         agerank, edurank, ins, smoke, yearrank, bmirank, wtmecadj,
FALSE         cluster, strata, drugdurcat) %>% filter(!is.na(bp_high),
FALSE         !is.na(bmirank), !is.na(drugdurcat)) %>% mutate(bp_high = as.integer(ifelse(bp_high ==
FALSE         "Yes", 1, 0))) %>% distinct(), nest = TRUE)
FALSE
FALSE Coefficients:
FALSE                          Estimate Std. Error t value Pr(>|t|)
FALSE (Intercept)             -4.89403    0.15054 -32.511  < 2e-16 ***
FALSE sexFemale               -0.14538    0.04385  -3.315 0.001293 **
FALSE raceHispanic            -0.23036    0.06114  -3.768 0.000284 ***
FALSE raceNon-Hispanic Black   0.51774    0.05108  10.137  < 2e-16 ***
FALSE raceNon-Hispanic Other   0.28094    0.08517   3.299 0.001363 **
FALSE agerank                  0.61374    0.01862  32.964  < 2e-16 ***
FALSE edurank                 -0.10299    0.02038  -5.054 2.06e-06 ***
FALSE insYes                   0.25049    0.06138   4.081 9.27e-05 ***
FALSE smokeCurrent             0.31455    0.04912   6.403 5.59e-09 ***
FALSE smokeFormer              0.19796    0.05280   3.749 0.000303 ***
FALSE yearrank                 0.00330    0.01629   0.203 0.839852
FALSE bmirank                  0.46966    0.01709  27.487  < 2e-16 ***
FALSE drugdurcat<1y           -0.06212    0.12246  -0.507 0.613119
FALSE drugdurcat>1y           -0.17636    0.45345  -0.389 0.698194
FALSE ---
FALSE Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
FALSE
FALSE (Dispersion parameter for binomial family taken to be 1.00329)
FALSE
FALSE Number of Fisher Scoring iterations: 5
```

Now we're looking at whether recreational drug use affects the frequency in which people reported high blood pressure, which it doesn't. Taking drugs within the past year (<1y) has no effect.

```
svy <- svydesign(
  id =   ~ cluster,
  strata =   ~ strata,
  weights =   ~ wtmecadj,
  data = users_everything %>%
    select(
      SEQN,
      dm,
      sex,
      race,
      agerank,
      edurank,
```

```
        ins,
        smoke,
        yearrank,
        bmirank,
        wtmecadj,
        cluster,
        strata,
        opioiddurcat
    ) %>%
    mutate(dm = case_when(dm == 'No' ~ 0, dm == 'Yes' ~ 1)) %>%
    filter(!is.na(dm), !is.na(bmirank), !is.na(opioiddurcat)) %>%
    distinct(),
  nest = TRUE
)
m <-
  svyglm(
    dm ~ sex + race + agerank + edurank + ins + smoke + yearrank + bmirank + opioiddurcat,
    design = svy,
    family = binomial(link = 'logit'),
    rescale = TRUE
  )
summary(m)
```

```
FALSE
FALSE Call:
FALSE svyglm(formula = dm ~ sex + race + agerank + edurank + ins +
FALSE     smoke + yearrank + bmirank + opioiddurcat, design = svy,
FALSE     family = binomial(link = "logit"), rescale = TRUE)
FALSE
FALSE Survey design:
FALSE svydesign(id = ~cluster, strata = ~strata, weights = ~wtmecadj,
FALSE     data = users_everything %>% select(SEQN, dm, sex, race, agerank,
FALSE         edurank, ins, smoke, yearrank, bmirank, wtmecadj, cluster,
FALSE         strata, opioiddurcat) %>% mutate(dm = case_when(dm ==
FALSE         "No" ~ 0, dm == "Yes" ~ 1)) %>% filter(!is.na(dm), !is.na(bmirank),
FALSE         !is.na(opioiddurcat)) %>% distinct(), nest = TRUE)
FALSE
FALSE Coefficients:
FALSE                          Estimate Std. Error t value Pr(>|t|)
FALSE (Intercept)             -6.923430   0.138581 -49.959  < 2e-16 ***
FALSE sexFemale               -0.289680   0.044310  -6.538 1.08e-09 ***
FALSE raceHispanic             0.561730   0.057250   9.812  < 2e-16 ***
FALSE raceNon-Hispanic Black   0.607122   0.050342  12.060  < 2e-16 ***
FALSE raceNon-Hispanic Other   0.917727   0.073288  12.522  < 2e-16 ***
FALSE agerank                  0.568174   0.012015  47.287  < 2e-16 ***
FALSE edurank                 -0.177737   0.022229  -7.996 4.35e-13 ***
FALSE insYes                   0.261859   0.062999   4.157 5.60e-05 ***
FALSE smokeCurrent             0.086085   0.054723   1.573   0.1179
FALSE smokeFormer              0.088876   0.044238   2.009   0.0465 *
FALSE yearrank                 0.043956   0.008662   5.075 1.21e-06 ***
FALSE bmirank                  0.548031   0.019276  28.431  < 2e-16 ***
FALSE opioiddurcatlong         0.473841   0.106131   4.465 1.64e-05 ***
FALSE opioiddurcatshort        0.115791   0.150771   0.768   0.4438
FALSE ---
```

```
FALSE Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
FALSE
FALSE (Dispersion parameter for binomial family taken to be 0.9237412)
FALSE
FALSE Number of Fisher Scoring iterations: 6
```

Here's the effect that long and short term opioid use has on presence of DM. Only long term opioid use impacts DM reporting, and the size of the effect is less than what we saw for chest pain.

```r
svy <- svydesign(
  id =   ~ cluster,
  strata =  ~ strata,
  weights =  ~ wtmecadj,
  data = users_everything %>%
    select(
      SEQN,
      dm,
      sex,
      race,
      agerank,
      edurank,
      ins,
      smoke,
      yearrank,
      bmirank,
      wtmecadj,
      cluster,
      strata,
      drugdurcat
    ) %>%
    mutate(dm = case_when(dm == 'No' ~ 0, dm == 'Yes' ~ 1)) %>%
    filter(!is.na(dm), !is.na(bmirank), !is.na(drugdurcat)) %>%
    distinct(),
  nest = TRUE
)
m <-
  svyglm(
    dm ~ sex + race + agerank + edurank + ins + smoke + yearrank + bmirank + drugdurcat,
    design = svy,
    family = binomial(link = 'logit'),
    rescale = TRUE
  )
summary(m)
```

```
FALSE
FALSE Call:
FALSE svyglm(formula = dm ~ sex + race + agerank + edurank + ins +
FALSE     smoke + yearrank + bmirank + drugdurcat, design = svy, family = binomial(link = "logit"),
FALSE     rescale = TRUE)
FALSE
FALSE Survey design:
FALSE svydesign(id = ~cluster, strata = ~strata, weights = ~wtmecadj,
FALSE     data = users_everything %>% select(SEQN, dm, sex, race, agerank,
FALSE         edurank, ins, smoke, yearrank, bmirank, wtmecadj, cluster,
FALSE         strata, drugdurcat) %>% mutate(dm = case_when(dm == "No" ~
```

```
FALSE          0, dm == "Yes" ~ 1)) %>% filter(!is.na(dm), !is.na(bmirank),
FALSE          !is.na(drugdurcat)) %>% distinct(), nest = TRUE)
FALSE
FALSE Coefficients:
FALSE                          Estimate Std. Error t value Pr(>|t|)
FALSE (Intercept)             -7.32824    0.21394 -34.254  < 2e-16 ***
FALSE sexFemale               -0.25264    0.07295  -3.463 0.000799 ***
FALSE raceHispanic             0.54919    0.08606   6.381 6.19e-09 ***
FALSE raceNon-Hispanic Black   0.55066    0.07950   6.926 4.91e-10 ***
FALSE raceNon-Hispanic Other   0.77986    0.10708   7.283 9.04e-11 ***
FALSE agerank                  0.68281    0.02290  29.820  < 2e-16 ***
FALSE edurank                 -0.20015    0.03626  -5.520 2.88e-07 ***
FALSE insYes                   0.35991    0.09033   3.984 0.000132 ***
FALSE smokeCurrent             0.13431    0.09259   1.451 0.150166
FALSE smokeFormer              0.14590    0.08935   1.633 0.105750
FALSE yearrank                 0.02810    0.02028   1.385 0.169160
FALSE bmirank                  0.56718    0.03008  18.854  < 2e-16 ***
FALSE drugdurcat<1y           -0.02844    0.19072  -0.149 0.881759
FALSE drugdurcat>1y           -0.31728    0.54784  -0.579 0.563848
FALSE ---
FALSE Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
FALSE
FALSE (Dispersion parameter for binomial family taken to be 0.968832)
FALSE
FALSE Number of Fisher Scoring iterations: 6
```

And here it shows that recreational drug use has no impact on DM reporting.

```r
svy <- svydesign(
  id =   ~ cluster,
  strata =   ~ strata,
  weights =   ~ wtmecadj,
  data = users_everything %>%
    select(
      SEQN,
      gen_health1,
      sex,
      race,
      agerank,
      edurank,
      ins,
      smoke,
      yearrank,
      bmirank,
      wtmecadj,
      cluster,
      strata,
      drugdurcat
    ) %>%
    filter(!is.na(gen_health1), !is.na(bmirank), !is.na(drugdurcat)) %>%
    mutate(gen_health1 = dense_rank(gen_health1)) %>%
    distinct(),
  nest = TRUE
)
m <-
```

```
  svyglm(
    gen_health1 ~ sex + race + agerank + edurank + ins + smoke + yearrank + bmirank + drugdurcat,
    design = svy,
    rescale = TRUE
  )
summary(m)
```

```
FALSE
FALSE Call:
FALSE svyglm(formula = gen_health1 ~ sex + race + agerank + edurank +
FALSE     ins + smoke + yearrank + bmirank + drugdurcat, design = svy,
FALSE     rescale = TRUE)
FALSE
FALSE Survey design:
FALSE svydesign(id = ~cluster, strata = ~strata, weights = ~wtmecadj,
FALSE     data = users_everything %>% select(SEQN, gen_health1, sex,
FALSE         race, agerank, edurank, ins, smoke, yearrank, bmirank,
FALSE         wtmecadj, cluster, strata, drugdurcat) %>% filter(!is.na(gen_health1),
FALSE         !is.na(bmirank), !is.na(drugdurcat)) %>% mutate(gen_health1 = dense_rank(gen_health1)) %>%
FALSE         distinct(), nest = TRUE)
FALSE
FALSE Coefficients:
FALSE                         Estimate Std. Error t value Pr(>|t|)
FALSE (Intercept)             3.864882   0.058158  66.455  < 2e-16 ***
FALSE sexFemale              -0.065611   0.014422  -4.549 1.57e-05 ***
FALSE raceHispanic           -0.293309   0.017480 -16.780  < 2e-16 ***
FALSE raceNon-Hispanic Black -0.187475   0.017466 -10.734  < 2e-16 ***
FALSE raceNon-Hispanic Other -0.234414   0.023512  -9.970  < 2e-16 ***
FALSE agerank                -0.043101   0.006049  -7.125 1.92e-10 ***
FALSE edurank                 0.191489   0.009238  20.728  < 2e-16 ***
FALSE insYes                  0.071684   0.019494   3.677 0.000389 ***
FALSE smokeCurrent           -0.328270   0.019328 -16.984  < 2e-16 ***
FALSE smokeFormer            -0.111619   0.020172  -5.533 2.72e-07 ***
FALSE yearrank               -0.018555   0.004194  -4.424 2.56e-05 ***
FALSE bmirank                -0.181588   0.005495 -33.045  < 2e-16 ***
FALSE drugdurcat<1y          -0.078281   0.040904  -1.914 0.058626 .
FALSE drugdurcat>1y          -0.046648   0.161115  -0.290 0.772799
FALSE ---
FALSE Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
FALSE
FALSE (Dispersion parameter for gaussian family taken to be 0.7044823)
FALSE
FALSE Number of Fisher Scoring iterations: 2
```

Nor does it affect a person's view of their general health.

```
svy <- svydesign(
  id =    ~ cluster,
  strata =   ~ strata,
  weights =   ~ wtmecadj,
  data = users_everything %>%
    select(
      SEQN,
      gen_health1,
      sex,
```

```
        race,
        agerank,
        edurank,
        ins,
        smoke,
        yearrank,
        bmirank,
        wtmecadj,
        cluster,
        strata,
        opioiddurcat
    ) %>%
    filter(!is.na(gen_health1), !is.na(bmirank), !is.na(opioiddurcat)) %>%
    mutate(gen_health1 = dense_rank(gen_health1)) %>%
    distinct(),
  nest = TRUE
)
m <-
  svyglm(
    gen_health1 ~ sex + race + agerank + edurank + ins + smoke + yearrank + bmirank + opioiddurcat,
    design = svy,
    rescale = TRUE
  )
summary(m)
```

```
FALSE
FALSE Call:
FALSE svyglm(formula = gen_health1 ~ sex + race + agerank + edurank +
FALSE     ins + smoke + yearrank + bmirank + opioiddurcat, design = svy,
FALSE     rescale = TRUE)
FALSE
FALSE Survey design:
FALSE svydesign(id = ~cluster, strata = ~strata, weights = ~wtmecadj,
FALSE     data = users_everything %>% select(SEQN, gen_health1, sex,
FALSE         race, agerank, edurank, ins, smoke, yearrank, bmirank,
FALSE         wtmecadj, cluster, strata, opioiddurcat) %>% filter(!is.na(gen_health1),
FALSE         !is.na(bmirank), !is.na(opioiddurcat)) %>% mutate(gen_health1 = dense_rank(gen_health1)) %
FALSE         distinct(), nest = TRUE)
FALSE
FALSE Coefficients:
FALSE                          Estimate Std. Error t value Pr(>|t|)
FALSE (Intercept)              3.787282   0.041056  92.247  < 2e-16 ***
FALSE sexFemale               -0.052682   0.011250  -4.683 7.21e-06 ***
FALSE raceHispanic            -0.309468   0.015937 -19.418  < 2e-16 ***
FALSE raceNon-Hispanic Black  -0.204921   0.014989 -13.671  < 2e-16 ***
FALSE raceNon-Hispanic Other  -0.253854   0.020013 -12.684  < 2e-16 ***
FALSE agerank                 -0.051440   0.003925 -13.106  < 2e-16 ***
FALSE edurank                  0.194488   0.006790  28.645  < 2e-16 ***
FALSE insYes                   0.104835   0.017580   5.963 2.33e-08 ***
FALSE smokeCurrent            -0.291087   0.013613 -21.383  < 2e-16 ***
FALSE smokeFormer             -0.085002   0.013149  -6.465 2.02e-09 ***
FALSE yearrank                -0.016041   0.002554  -6.280 5.03e-09 ***
FALSE bmirank                 -0.164042   0.004565 -35.931  < 2e-16 ***
FALSE opioiddurcatlong        -0.612283   0.031323 -19.548  < 2e-16 ***
```

```
FALSE opioiddurcatshort        -0.231193   0.042519  -5.437 2.70e-07 ***
FALSE ---
FALSE Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
FALSE
FALSE (Dispersion parameter for gaussian family taken to be 0.7272799)
FALSE
FALSE Number of Fisher Scoring iterations: 2
```

But opioids do, in that opioid is inversely associated with a person's view of their general health, so if one uses opioids, they're more likely to have a more negative view of their health, and the degree of the relationship is greater with a longer duration of opioid use.

```
plot_vars('opioiddurcat',
          'cp',
          data = users_everything,
          vars = c('ins', 'sex')) +
  th +
  labs(y = '',
       x = 'Duration Using Opioids',
       title = 'Sex vs Insurance',
       fill = 'CP')
```
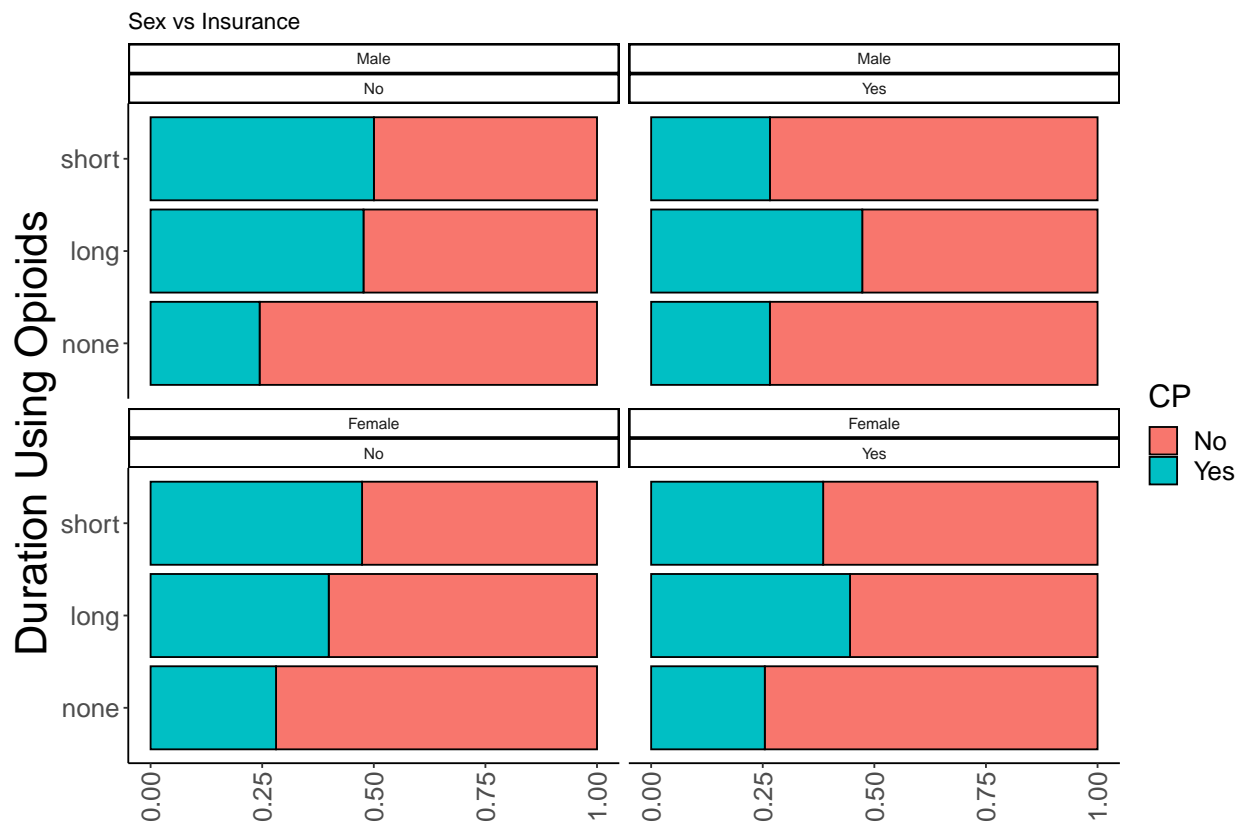
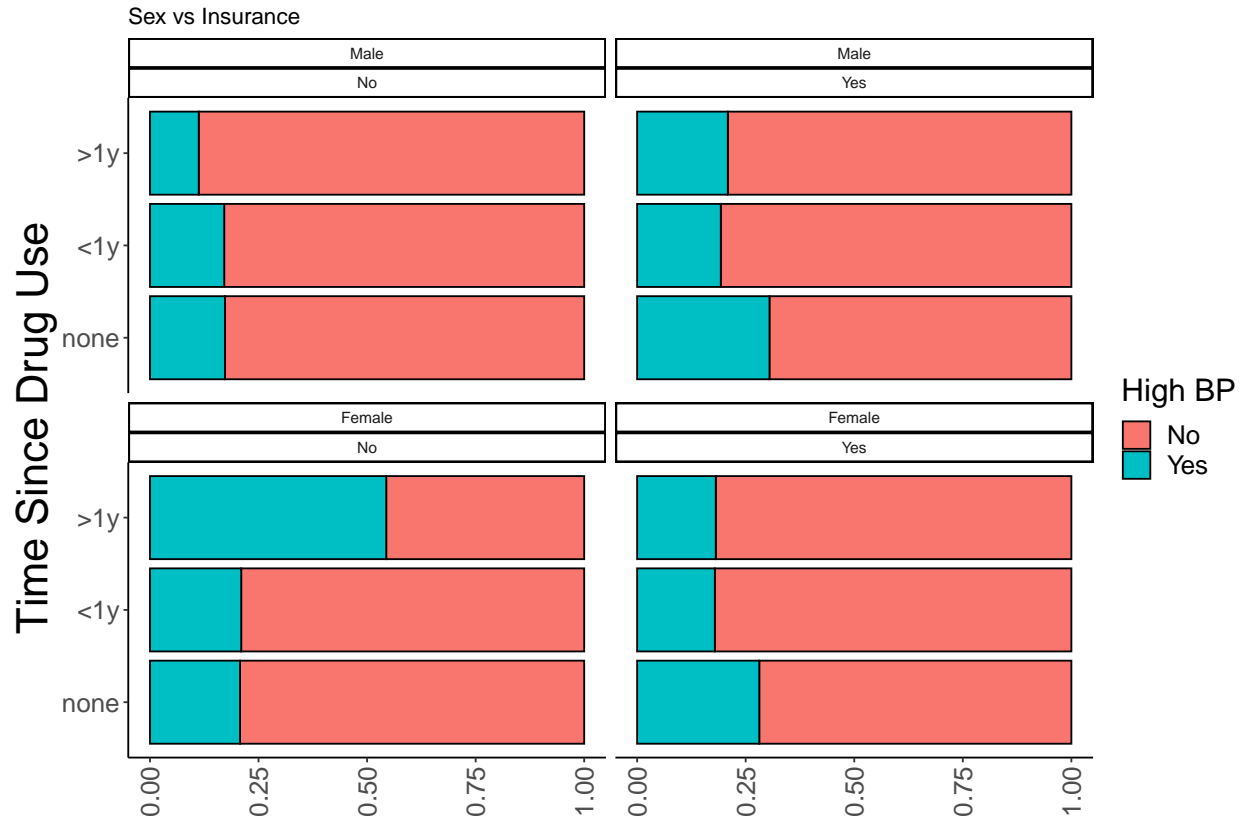

```
plot_vars('opioiddurcat',
          'bp_high',
          data = users_everything,
          vars = c('ins', 'sex')) +
  th +
```
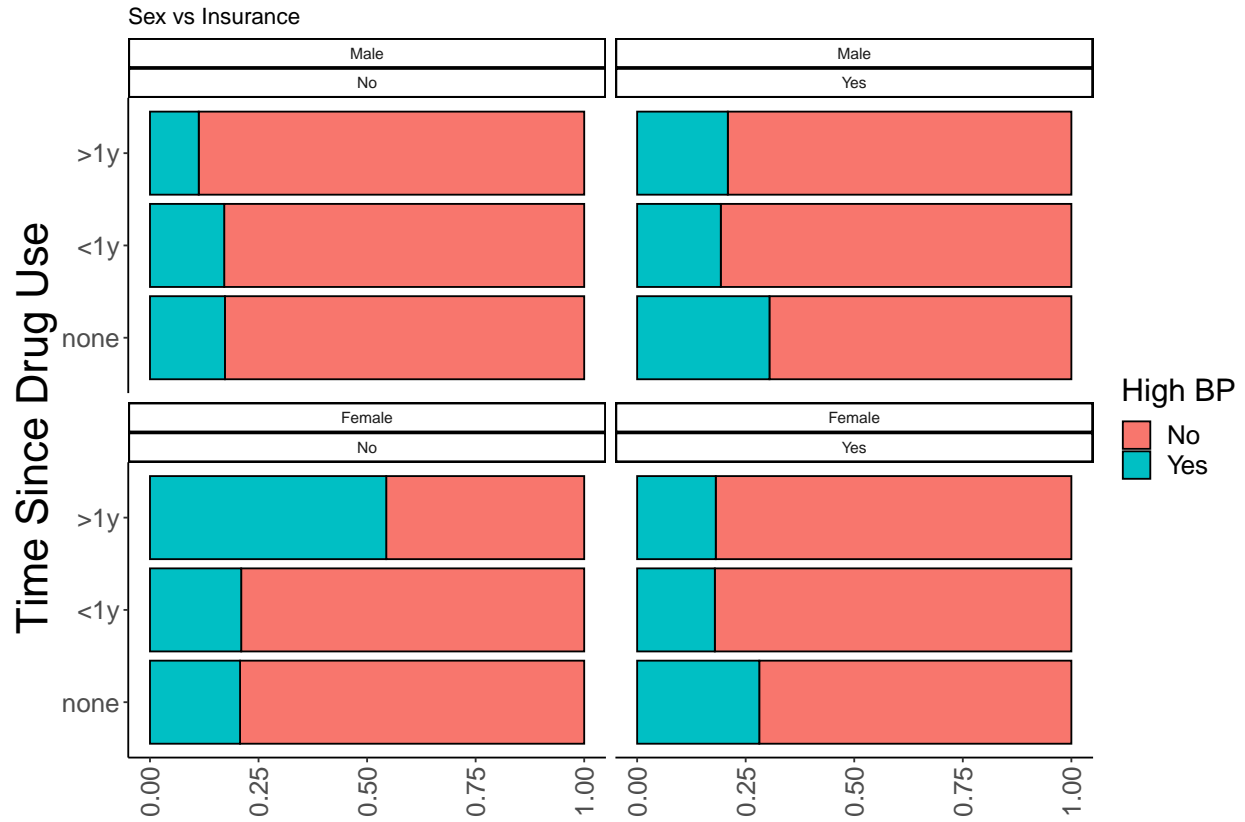
```
labs(y = '',
     x = 'Duration Using Opioids',
     title = 'Sex vs Insurance',
     fill = 'High BP')
```


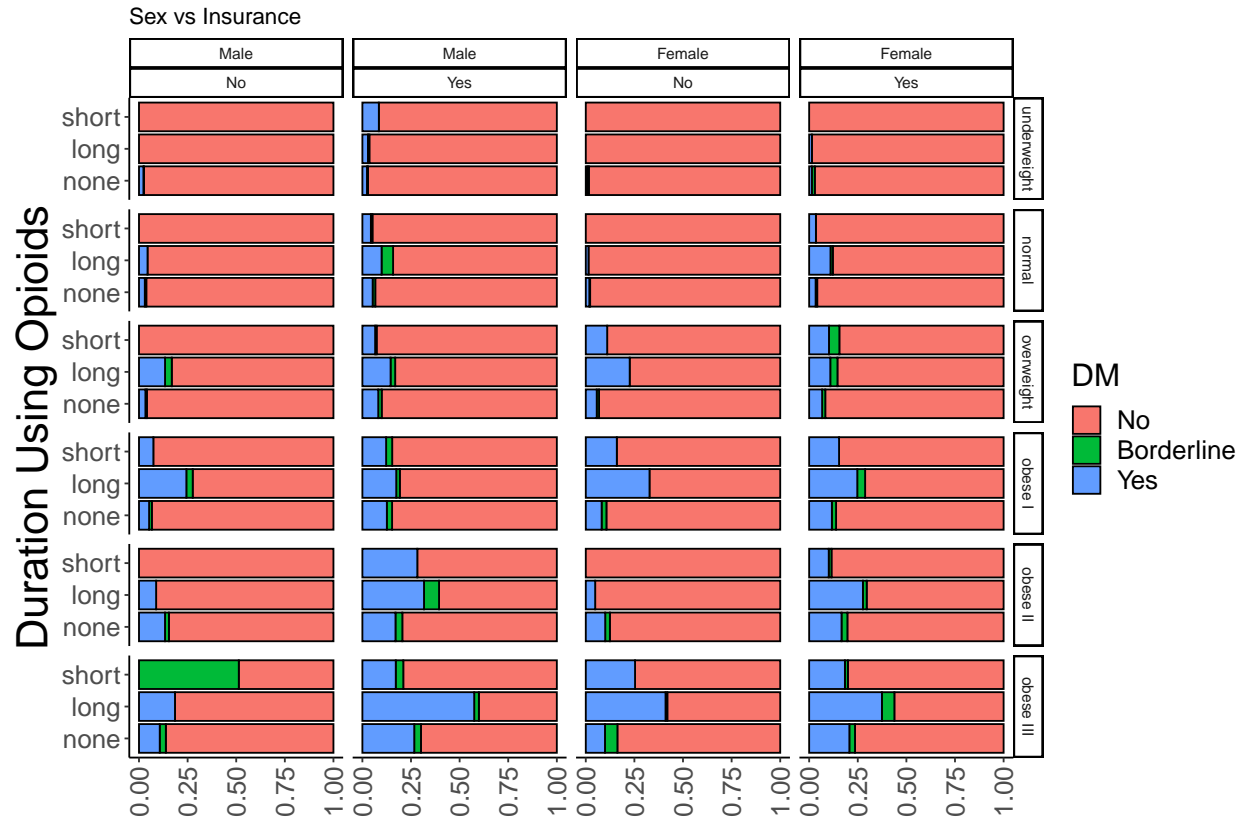Sex vs Insurance

```
plot_vars('drugdurcat',
          'bp_high',
          data = users_everything,
          vars = c('ins', 'sex')) +
  th +
  labs(y = '',
       x = 'Time Since Drug Use',
       title = 'Sex vs Insurance',
       fill = 'High BP')
```

Sex vs Insurance

```
plot_vars('drugdurcat',
          'bp_high',
          data = users_everything,
          vars = c('ins', 'sex')) +
  th +
  labs(y = '',
       x = 'Time Since Drug Use',
       title = 'Sex vs Insurance',
       fill = 'High BP')
```
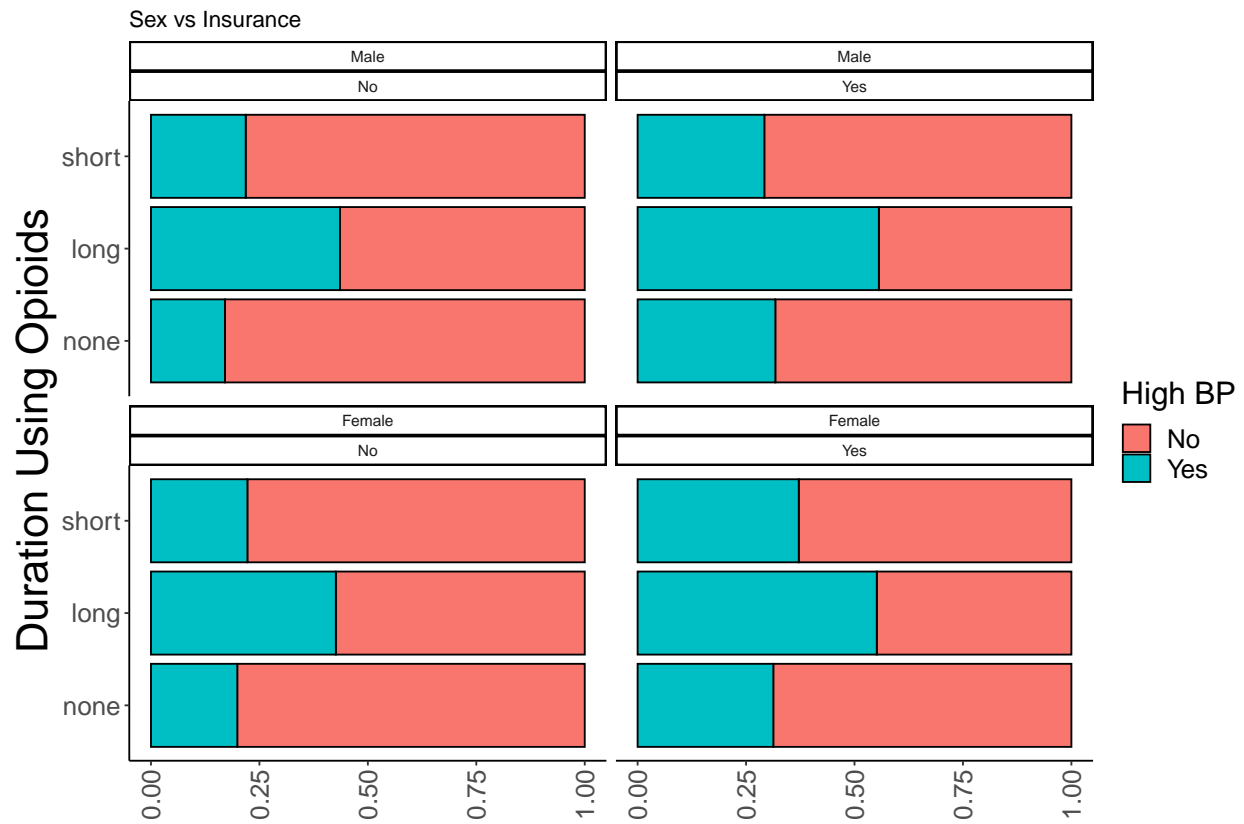
Sex vs Insurance

```
plot_vars('opioiddurcat',
          'dm',
          data = users_everything,
          vars = c('ins', 'sex', 'bmicat')) +
  th +
  labs(y = '',
       x = 'Duration Using Opioids',
       title = 'Sex vs Insurance',
       fill = 'DM')
```

```
plot_vars('drugdurcat',
          'dm',
          data = users_everything,
          vars = c('ins', 'sex', 'bmicat')) +
  th +
  labs(y = '',
       x = 'Time Since Drug Use',
       title = 'Sex vs Insurance',
       fill = 'DM')
```

Sex vs Insurance