

Recreate Paper

SW

7/28/2020

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

Functions

```
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
  }

  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))
}

# druguse <- get_data('RXQ_RX', 'Q')
# druginfo <- get_data('RXQ_DRUG', 'Q')
# demo <- get_data('DEMO', 'DEMO')
# demo_raw <- get_data('DEMO', 'DEMO', translate=FALSE)
# body <- get_data('BMX', 'EXAM')

```

```

# preg <- get_data('RHQ','Q') #RHD143
# cancer <- get_data('MCQ','Q') #MCQ220 starts_with(MCQ240)
# smoke <- get_data('SMQ','Q')
# ins <- get_data('HIQ','Q')
# otherdruguse <- get_data('DUQ','Q')

# saveRDS(list(druguse=druguse, druginfo=druginfo,
#             otherdrug=otherdruguse, demo=demo,
#             demo_raw=demo_raw, body=body, preg=preg, cancer=cancer,
#             smoke=smoke, ins=ins),
#         'D:/Downloads/substance_use_analysis/tifdata_paper.rds')

```

Load and Prepare Data

The data is loaded and prepared based on the paper's description.

```

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

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))

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)

paper <- dat$druguse %>%
  select(SEQN, year, druguse=RXDUSE, drug=RXDDRUG, drugid=RXDDRGID,
         drugdur=RXDDAYS) %>%
  filter(druguse %in% c('Yes', 'No')) %>%
  mutate(druguse=ifelse(drugid %in% opioids$id, 'Yes', 'No'),
         drugdur=ifelse(druguse=='Yes', as.integer(drugdur), NA),
         drugdurcat=case_when(druguse == 'No' ~ 'none',
                              drugdur < 90 ~ 'short',
                              drugdur >= 90 & drugdur <= 25550 ~ 'long')) %>%
  filter(!is.na(drugdurcat)) %>%

```

```

filter(year %in% c(2003,2005,2007,2009,2011,2013,2015)) %>%
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 >= 35 & age <= 79) %>%
left_join(dat$body %>% select(SEQN,year,bmi=BMXBMI),by=c('SEQN','year')) %>%
filter(bmi > 15 & bmi < 80) %>%
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 >= 35 & age < 45 ~ '35-44',

```

```

        age >= 45 & age < 55 ~ '45-54',
        age >= 55 & age < 65 ~ '55-64',
        age >= 65 & age <= 79 ~ '65-79')) %>%
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='Female'),
        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='Less than high school'),
        ins=relevel(as.factor(ins),ref='No'),
        smoke=relevel(as.factor(smoke),ref='Never'),
        year=relevel(as.factor(year),ref='2003')) %>%
mutate(wtint7yr=wtint2yr/7,wtmec7yr=wtmec2yr/7,
        strata=dense_rank(strata),cluster=dense_rank(cluster),
        drugdurrank=case_when(drugdurcat == 'none' ~ 1,
                                drugdurcat == 'short' ~ 2,
                                drugdurcat == 'long' ~ 3)) %>%
select(SEQN,bmicat,agecat,sex,race,edu,ins,smoke,year,strata,cluster,
        drugdurcat,drugdurrank,
        druguse,
        wtint7yr,wtmec7yr) %>% distinct()

```

Fit Multinomial Regression Model

We fit a multinomial regression model and extract the marginal effects.

```

# mod2 <- multinom(drugdurcat ~ bmicat + agecat + sex + race + edu + ins +
#                   smoke + year,
#                   weights=wtmec7yr,
#                   data=paper)
# mod2 <- multinom(drugdurcat ~ bmicat*agecat + sex*race + edu + ins +
#                   smoke + year,
#                   # weights=wtmec7yr,
#                   data=paper)
mod2 <- multinom(drugdurcat ~ bmicat + agecat + sex + race + edu + ins +
                  smoke + year,
                  # weights=wtmec7yr,
                  data=paper)
out <- Effect(c('bmicat'),mod2)

```

Plot Results

```
as_tibble(out$prob) %>%
  bind_cols(bmicat=out$x) %>%
  gather(drugdur,prob,-bmicat) %>%
  mutate(drugdur=str_replace(drugdur,'prob\\.','')) %>%
  left_join(as_tibble(out$lower.prob) %>%
    bind_cols(bmicat=out$x) %>%
    gather(drugdur,lower,-bmicat) %>%
    mutate(drugdur=str_replace(drugdur,'L\\.prob\\.',''),
      by=c('bmicat','drugdur')) %>%
  left_join(as_tibble(out$upper.prob) %>%
    bind_cols(bmicat=out$x) %>%
    gather(drugdur,upper,-bmicat) %>%
    mutate(drugdur=str_replace(drugdur,'U\\.prob\\.',''),
      by=c('bmicat','drugdur')) %>%
  mutate(bmicat=factor(bmicat,ordered=TRUE,
    levels=c('underweight','normal','overweight','obese I',
      'obese II','obese III'))) %>%
  filter(drugdur %in% c('short','long')) %>%
  ggplot(aes(x=bmicat,y=prob,ymin=lower,ymax=upper,
    color=drugdur,group=drugdur)) +
  geom_line() +
  geom_linerange() +
  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=45,hjust=1),
    axis.title=element_text(size=25)) +
  labs(color='Drug Duration',x='BMI',y='Probability')
```

