R Analysis Example Replication C12

```
# Chapter 12 Multiple Imputation

library(survey)
library(mice)

# Read in Nhanes C11 data set, set up for missing data imputation and analysis 2
nhanesc12  <- read.table(file = "P:/ASDA 2/Data sets/nhanes 2011_2012/c12_impute_subset_nhanes1112.csv", sep =
",", header = T, as.is=T)
nhanesc12$marcat=factor(nhanesc12$marcat)

# subset to just records used in analysis
nhanesc12_sub <- nhanesc12[ which(nhanesc12$age18p==1 & nhanesc12$wtmec2yr >0),]
summary(nhanesc12_sub)

# Missing data pattern
md.pattern(nhanesc12_sub)

# create an indicator of high blood pressure
nhanesc12_sub$high_diastolic <- ifelse(nhanesc12_sub$bpxdi1_1 >=90, 1,0)
summary(nhanesc12_sub$high_diastolic)

# survey design
nhanessvy <- svydesign(strata=~sdmvstra, id=~sdmvpsu, weights=~wtmec2yr, data=nhanesc12_sub, nest=T)

# Complete Case Analysis
# Obtain means for 3 continuous variables imputed using method 1 as well, Table 12.3
show(ex12_3a <- svymean(~bmxbmi, nhanessvy, se=T, na.rm=T, ci=T))
show(ex12_3b <- svymean(~bpxdi1_1, nhanessvy, se=T, na.rm=T, ci=T))
show(ex12_3c <- svymean(~indfmpir, nhanessvy, se=T, na.rm=T, ci=T))
# High Blood Pressure Mean, Complete Case, Table 12.4
(ex12_4 <- svymean(~factor(high_diastolic),nhanessvy, se=T, na.rm=T, deff=T, ci=T, keep.vars=T))
confint(ex12_4)
# Logistic Regression Complete Case Analysis with Design Correction, Table 12.5
mod12_5 <- svyglm(high_diastolic ~ factor(ridreth1) + factor(riagendr)+ agec + agecsq,family=quasibinomial,
design=nhanessvy)
summary(mod12_5)
```

```
# Method 1 with Design Variables in Model

nhanesc12_subm1 <- nhanesc12[ which(nhanesc12$age18p==1 & nhanesc12$wtmec2yr >0),]
#nhanesc12_subm1$descode=factor(nhanesc12_subm1$descode)
summary(nhanesc12_subm1)

# use mice to impute and specify type of default method for imputation models

# run without custom predictor matrix first
impm1 <- mice(nhanesc12_subm1, n.imp=5, seed=2016, defaultMethod=c("norm","logreg","polyreg"))
impm1$predictorMatrix
summary(impm1)

# add a predictor matrix to control imputation model predictors for each imputed variable
pred <- impm1$predictorMatrix
pred[,"sdmvpsu"] <- 0
pred[,"sdmvstra"] <- 0
pred[,"seqn"] <- 0
pred[,"age18p"] <- 0
pred[,"descode"] <- 1
impm1$predictorMatrix
pred
impm1_pred <- mice(nhanesc12_subm1, pred=pred, n.imp=5, seed=2016, defaultMethod=c("norm","logreg","polyreg"))
summary(impm1_pred)

# create new high diastolic blood pressure base on imputed bpxdi1_1, prepare long data set and then back to mids
after computation
longm1 <- complete(impm1_pred, action='long', include=TRUE)
longm1$high_diastolic <- ifelse(longm1$bpxdi1_1 >=90,1,0)
summary(longm1)

#use as.mids() to convert back to mids object
impm1a  <- as.mids(longm1)

# convert mids to data useable for work in mitools
library(mitools)
mydatam1 <- imputationList(lapply(1:5, complete, x=impm1a))
summary(mydatam1)

# set survey design
library(survey)
desm1 <- svydesign(id=~sdmvstra, strat=~sdmvpsu, weight=~wtmec2yr, data=(mydatam1), nest=TRUE)
summary(desm1)

# run design based logistic model with svyglm using 5 imputed data sets contained in desm1, Taable 12.5
fitm1 <- with(desm1, svyglm (high_diastolic ~ factor(ridreth1) + factor(riagendr) + agec + agecsq,
family=quasibinomial))
summary(MIcombine(fitm1))

# combined mean high blood pressure with design adjustment, Table 12.4
fitm1_mean <- with(desm1, svymean(~factor(high_diastolic), se=T, na.rm=T, ci=T ))
# mean high blood pressure for each imputed data set
fitm1_mean
```

```
# Use MIcombine for overall combined and design-adjusted mean/se
summary(MIcombine(fitm1_mean))
# Obtain means for 3 continuous variables imputed using method 1 as well, Table 12.3
fitm1_ex12_3 <- with(desm1, svymean(~bmxbmi+bpxdi1_1+indfmpir), se=T, na.rm=T, ci=T )
fitm1_ex12_3
```

```
# Method 2 WITHOUT Design Variables in Model
# return to original data without the CC high diastolic blood pressure
nhanesc12_subm2 <- nhanesc12[ which(nhanesc12$age18p==1 & nhanesc12$wtmec2yr >0),]
#nhanesc12_subm2$descode=factor(nhanesc12_subm2$descode)
summary(nhanesc12_subm2)

# use mice to impute and specify type of default method for imputation models

# run without custom predictor matrix first
impm2 <- mice(nhanesc12_subm2, n.imp=5, seed=2016, defaultMethod=c("norm","logreg","polyreg"))
impm2$predictorMatrix
summary(impm2)

# add a predictor matrix to control imputation model predictors for each imputed variable
pred <- impm2$predictorMatrix
pred[,"sdmvpsu"] <- 0
pred[,"sdmvstra"] <- 0
pred[,"seqn"] <- 0
pred[,"age18p"] <- 0
pred[,"descode"] <- 0
pred[,"wtmec2yr"] <- 0
impm2$predictorMatrix
pred
impm2_pred <- mice(nhanesc12_subm2, pred=pred, n.imp=5, seed=2016, defaultMethod=c("norm","logreg","polyreg"))
summary(impm2_pred)

# create new high diastolic blood pressure base on imputed bpxdi1_1, prepare long data set and then back to mids
after computation
longm2 <- complete(impm2_pred, action='long', include=TRUE)
longm2$high_diastolic <- ifelse(longm2$bpxdi1_1 >=90,1,0)
summary(longm2)

#use as.mids() to convert back to mids object
impm2a  <- as.mids(longm2)

# convert mids to data useable for work in mitools
library(mitools)
mydatam2 <- imputationList(lapply(1:5, complete, x=impm2a))
summary(mydatam2)

# set survey design
library(survey)
desm2 <- svydesign(id=~sdmvstra, strat=~sdmvpsu, weight=~wtmec2yr, data=(mydatam2), nest=TRUE)
summary(desm2)

# run design based logistic model with svyglm using 5 imputed data sets contained in desm1, Taable 12.5
fitm2 <- with(desm2, svyglm (high_diastolic ~ factor(ridreth1) + factor(riagendr) + agec + agecsq,
family=quasibinomial))
summary(MIcombine(fitm2))

# combined mean high blood pressure with design adjustment, Table 12.4
fitm2_mean <- with(desm2, svymean(~factor(high_diastolic), se=T, na.rm=T, ci=T ))
```

```
# mean high blood pressure for each imputed data set
fitm2_mean
# Use MIcombine for overall combined and design-adjusted mean/se
summary(MIcombine(fitm2_mean))
```

 # Note: FEFI method available in R as of early June 2017, see https://sites.google.com/view/jaekwangkim/software for more information, this will be included on ASDA website in the near future.

Output R Analysis Example Replication C12

```
> # Chapter 12 Multiple Imputation
> # stata code as guide
> library(survey)
> library(mice)
>
> # Read in Nhanes C11 data set, set up for missing data imputation and analysis 2
> nhanesc12  <- read.table(file = "P:/ASDA 2/Data sets/nhanes 2011_2012/c12_impute_subset_nhanes1112.csv", sep =
",", header = T, as.is=T)
> nhanesc12$marcat=factor(nhanesc12$marcat)
>
> # subset to just records used in analysis
> nhanesc12_sub <- nhanesc12[ which(nhanesc12$age18p==1 & nhanesc12$wtmec2yr >0),]
> summary(nhanesc12_sub)
      seqn          riagendr        ridreth1        wtmec2yr         sdmvpsu         sdmvstra
 Min.   :62161   Min.   :1.000   Min.   :1.000   Min.   :  4413   Min.   :1.000   Min.   : 90.00
 1st Qu.:64611   1st Qu.:1.000   1st Qu.:3.000   1st Qu.: 16174   1st Qu.:1.000   1st Qu.: 92.00
 Median :67109   Median :2.000   Median :3.000   Median : 24567   Median :2.000   Median : 96.00
 Mean   :67076   Mean   :1.506   Mean   :3.301   Mean   : 41318   Mean   :1.638   Mean   : 95.87
 3rd Qu.:69533   3rd Qu.:2.000   3rd Qu.:4.000   3rd Qu.: 45238   3rd Qu.:2.000   3rd Qu.: 99.00
 Max.   :71915   Max.   :2.000   Max.   :5.000   Max.   :222580   Max.   :3.000   Max.   :103.00

    indfmpir         bmxbmi          age18p       age          marcat        descode          bpxdi1_1
 Min.   :0.000   Min.   :13.40   Min.   :1   Min.   :18.00   1   :2991   Min.   : 901.0   Min.   : 10.00
 1st Qu.:0.950   1st Qu.:23.80   1st Qu.:1   1st Qu.:31.00   2   :1183   1st Qu.: 922.0   1st Qu.: 64.00
 Median :1.840   Median :27.40   Median :1   Median :47.00   3   :1141   Median : 961.0   Median : 72.00
 Mean   :2.374   Mean   :28.62   Mean   :1   Mean   :47.17   NA's: 300   Mean   : 960.3   Mean   : 71.02
 3rd Qu.:3.933   3rd Qu.:32.00   3rd Qu.:1   3rd Qu.:62.00               3rd Qu.: 992.0   3rd Qu.: 78.00
 Max.   :5.000   Max.   :82.10   Max.   :1   Max.   :80.00               Max.   :1032.0   Max.   :120.00
 NA's   :487     NA's   :90                                                             NA's   :503
      agec             agecsq             age4cat
 Min.   :-28.3552   Min.   :   0.1261   Min.   :1.000
 1st Qu.:-15.3552   1st Qu.:  58.4436   1st Qu.:2.000
 Median :  0.6448   Median : 244.7610   Median :3.000
 Mean   :  0.8099   Mean   : 345.0635   Mean   :2.594
 3rd Qu.: 15.6448   3rd Qu.: 559.0785   3rd Qu.:3.000
 Max.   : 33.6448   Max.   :1131.9752   Max.   :4.000

>
> # Missing data pattern
> md.pattern(nhanesc12_sub)
     seqn riagendr ridreth1 wtmec2yr sdmvpsu sdmvstra age18p age descode agec agecsq age4cat bmxbmi marcat
4416    1        1        1        1       1        1      1   1       1    1      1       1      1      1
 369    1        1        1        1       1        1      1   1       1    1      1       1      1      1
  48    1        1        1        1       1        1      1   1       1    1      1       1      0      1
 230    1        1        1        1       1        1      1   1       1    1      1       1      1      0
 386    1        1        1        1       1        1      1   1       1    1      1       1      1      1
  12    1        1        1        1       1        1      1   1       1    1      1       1      0      1
  31    1        1        1        1       1        1      1   1       1    1      1       1      1      0
   6    1        1        1        1       1        1      1   1       1    1      1       1      0      0
  62    1        1        1        1       1        1      1   1       1    1      1       1      1      1
  18    1        1        1        1       1        1      1   1       1    1      1       1      0      1
```

```
  22    1        1        1        1       1        1      1   1       1    1      1       1      1      0
   4    1        1        1        1       1        1      1   1       1    1      1       1      0      1
   9    1        1        1        1       1        1      1   1       1    1      1       1      1      0
   2    1        1        1        1       1        1      1   1       1    1      1       1      0      0
        0        0        0        0       0        0      0   0       0    0      0       0     90    300
     indfmpir bpxdi1_1
4416        1        1    0
 369        0        1    1
  48        1        1    1
 230        1        1    1
 386        1        0    1
  12        0        1    2
  31        0        1    2
   6        1        1    2
  62        0        0    2
  18        1        0    2
  22        1        0    2
   4        0        0    3
   9        0        0    3
   2        1        0    3
         487      503 1380
>
> # create an indicator of high blood pressure
> nhanesc12_sub$high_diastolic <- ifelse(nhanesc12_sub$bpxdi1_1 >=90, 1,0)
> summary(nhanesc12_sub$high_diastolic)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
  0.000   0.000   0.000   0.062   0.000   1.000     503
>
> # survey design
> nhanessvy <- svydesign(strata=~sdmvstra, id=~sdmvpsu, weights=~wtmec2yr, data=nhanesc12_sub, nest=T)
>
> # Complete Case Analysis
> # Obtain means for 3 continuous variables imputed using method 1 as well, Table 12.3
> show(ex12_3a <- svymean(~bmxbmi, nhanessvy, se=T, na.rm=T, ci=T))
         mean    SE
bmxbmi 28.623 0.214
> show(ex12_3b <- svymean(~bpxdi1_1, nhanessvy, se=T, na.rm=T, ci=T))
           mean     SE
bpxdi1_1 71.609 0.5047
> show(ex12_3c <- svymean(~indfmpir, nhanessvy, se=T, na.rm=T, ci=T))
           mean     SE
indfmpir 2.8592 0.1064
>
>
> # High Blood Pressure Mean, Complete Case, Table 12.4
> (ex12_4 <- svymean(~factor(high_diastolic),nhanessvy, se=T, na.rm=T, deff=T, ci=T, keep.vars=T))
                             mean        SE   DEff
factor(high_diastolic)0 0.9391787 0.0079624 5.6729
factor(high_diastolic)1 0.0608213 0.0079624 5.6729
> confint(ex12_4)
                             2.5 %     97.5 %
factor(high_diastolic)0 0.92357258 0.95478473
factor(high_diastolic)1 0.04521527 0.07642742
```

```
>
> # Logistic Regression Complete Case Analysis with Design Correction, Table 12.5
> mod12_5 <- svyglm(high_diastolic ~ factor(ridreth1) + factor(riagendr)+ agec + agecsq,family=quasibinomial,
design=nhanessvy)
> summary(mod12_5)

Call:
svyglm(formula = high_diastolic ~ factor(ridreth1) + factor(riagendr) +
    agec + agecsq, family = quasibinomial, design = nhanessvy)

Survey design:
svydesign(strata = ~sdmvstra, id = ~sdmvpsu, weights = ~wtmec2yr,
    data = nhanesc12_sub, nest = T)

Coefficients:
                     Estimate Std. Error t value Pr(>|t|)
(Intercept)       -2.2498852  0.1985021 -11.334 4.99e-07 ***
factor(ridreth1)2 -0.7256815  0.2449515  -2.963 0.014226 *
factor(ridreth1)3  0.1312534  0.2245498   0.585 0.571822
factor(ridreth1)4  0.6582416  0.2463534   2.672 0.023414 *
factor(ridreth1)5  0.0498803  0.2447868   0.204 0.842620
factor(riagendr)2 -0.5467630  0.2077358  -2.632 0.025075 *
agec               0.0084599  0.0069603   1.215 0.252108
agecsq            -0.0016157  0.0002767  -5.838 0.000164 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasibinomial family taken to be 1.019763)

Number of Fisher Scoring iterations: 6
```

```
> # Method 1 with Design Variables in Model
> # return to original data without the CC high diastolic blood pressure
> nhanesc12_subm1 <- nhanesc12[ which(nhanesc12$age18p==1 & nhanesc12$wtmec2yr >0),]
> #nhanesc12_subm1$descode=factor(nhanesc12_subm1$descode)
>
> summary(nhanesc12_subm1)
      seqn          riagendr        ridreth1        wtmec2yr         sdmvpsu         sdmvstra
 Min.   :62161   Min.   :1.000   Min.   :1.000   Min.   :  4413   Min.   :1.000   Min.   : 90.00
 1st Qu.:64611   1st Qu.:1.000   1st Qu.:3.000   1st Qu.: 16174   1st Qu.:1.000   1st Qu.: 92.00
 Median :67109   Median :2.000   Median :3.000   Median : 24567   Median :2.000   Median : 96.00
 Mean   :67076   Mean   :1.506   Mean   :3.301   Mean   : 41318   Mean   :1.638   Mean   : 95.87
 3rd Qu.:69533   3rd Qu.:2.000   3rd Qu.:4.000   3rd Qu.: 45238   3rd Qu.:2.000   3rd Qu.: 99.00
 Max.   :71915   Max.   :2.000   Max.   :5.000   Max.   :222580   Max.   :3.000   Max.   :103.00

    indfmpir         bmxbmi          age18p       age         marcat        descode          bpxdi1_1
 Min.   :0.000   Min.   :13.40   Min.   :1   Min.   :18.00   1   :2991   Min.   : 901.0   Min.   : 10.00
 1st Qu.:0.950   1st Qu.:23.80   1st Qu.:1   1st Qu.:31.00   2   :1183   1st Qu.: 922.0   1st Qu.: 64.00
 Median :1.840   Median :27.40   Median :1   Median :47.00   3   :1141   Median : 961.0   Median : 72.00
 Mean   :2.374   Mean   :28.62   Mean   :1   Mean   :47.17   NA's: 300   Mean   : 960.3   Mean   : 71.02
 3rd Qu.:3.933   3rd Qu.:32.00   3rd Qu.:1   3rd Qu.:62.00               3rd Qu.: 992.0   3rd Qu.: 78.00
 Max.   :5.000   Max.   :82.10   Max.   :1   Max.   :80.00               Max.   :1032.0   Max.   :120.00
 NA's   :487     NA's   :90                                                             NA's   :503
      agec             agecsq            age4cat
 Min.   :-28.3552   Min.   :   0.1261   Min.   :1.000
 1st Qu.:-15.3552   1st Qu.:  58.4436   1st Qu.:2.000
 Median :  0.6448   Median : 244.7610   Median :3.000
 Mean   :  0.8099   Mean   : 345.0635   Mean   :2.594
 3rd Qu.: 15.6448   3rd Qu.: 559.0785   3rd Qu.:3.000
 Max.   : 33.6448   Max.   :1131.9752   Max.   :4.000


> # use mice to impute and specify type of default method for imputation models

> # run without custom predictor matrix first
> impm1 <- mice(nhanesc12_subm1, n.imp=5, seed=2016, defaultMethod=c("norm","logreg","polyreg"))

 iter imp variable
  1   1  indfmpir  bmxbmi  marcat  bpxdi1_1
  1   2  indfmpir  bmxbmi  marcat  bpxdi1_1
  1   3  indfmpir  bmxbmi  marcat  bpxdi1_1
  1   4  indfmpir  bmxbmi  marcat  bpxdi1_1
  1   5  indfmpir  bmxbmi  marcat  bpxdi1_1
  2   1  indfmpir  bmxbmi  marcat  bpxdi1_1
  2   2  indfmpir  bmxbmi  marcat  bpxdi1_1
  2   3  indfmpir  bmxbmi  marcat  bpxdi1_1
  2   4  indfmpir  bmxbmi  marcat  bpxdi1_1
  2   5  indfmpir  bmxbmi  marcat  bpxdi1_1
  3   1  indfmpir  bmxbmi  marcat  bpxdi1_1
  3   2  indfmpir  bmxbmi  marcat  bpxdi1_1
  3   3  indfmpir  bmxbmi  marcat  bpxdi1_1
  3   4  indfmpir  bmxbmi  marcat  bpxdi1_1
  3   5  indfmpir  bmxbmi  marcat  bpxdi1_1
  4   1  indfmpir  bmxbmi  marcat  bpxdi1_1
```

```
  4   2  indfmpir  bmxbmi  marcat  bpxdi1_1
  4   3  indfmpir  bmxbmi  marcat  bpxdi1_1
  4   4  indfmpir  bmxbmi  marcat  bpxdi1_1
  4   5  indfmpir  bmxbmi  marcat  bpxdi1_1
  5   1  indfmpir  bmxbmi  marcat  bpxdi1_1
  5   2  indfmpir  bmxbmi  marcat  bpxdi1_1
  5   3  indfmpir  bmxbmi  marcat  bpxdi1_1
  5   4  indfmpir  bmxbmi  marcat  bpxdi1_1
  5   5  indfmpir  bmxbmi  marcat  bpxdi1_1
> impm1$predictorMatrix
         seqn riagendr ridreth1 wtmec2yr sdmvpsu sdmvstra indfmpir bmxbmi age18p age marcat descode
seqn        0        0        0        0       0        0        0      0      0   0      0        0
riagendr    0        0        0        0       0        0        0      0      0   0      0        0
ridreth1    0        0        0        0       0        0        0      0      0   0      0        0
wtmec2yr    0        0        0        0       0        0        0      0      0   0      0        0
sdmvpsu     0        0        0        0       0        0        0      0      0   0      0        0
sdmvstra    0        0        0        0       0        0        0      0      0   0      0        0
indfmpir    1        1        1        1       1        1        0      1      0   1      1        0
bmxbmi      1        1        1        1       1        1        1      0      0   1      1        0
age18p      0        0        0        0       0        0        0      0      0   0      0        0
age         0        0        0        0       0        0        0      0      0   0      0        0
marcat      1        1        1        1       1        1        1      1      0   1      0        0
descode     0        0        0        0       0        0        0      0      0   0      0        0
bpxdi1_1    1        1        1        1       1        1        1      1      0   1      1        0
agec        0        0        0        0       0        0        0      0      0   0      0        0
agecsq      0        0        0        0       0        0        0      0      0   0      0        0
age4cat     0        0        0        0       0        0        0      0      0   0      0        0
         bpxdi1_1 agec agecsq age4cat
seqn            0    0      0       0
riagendr        0    0      0       0
ridreth1        0    0      0       0
wtmec2yr        0    0      0       0
sdmvpsu         0    0      0       0
sdmvstra        0    0      0       0
indfmpir        1    0      1       1
bmxbmi          1    0      1       1
age18p          0    0      0       0
age             0    0      0       0
marcat          1    0      1       1
descode         0    0      0       0
bpxdi1_1        0    0      1       1
agec            0    0      0       0
agecsq          0    0      0       0
age4cat         0    0      0       0
>
> summary(impm1)
Multiply imputed data set
Call:
mice(data = nhanesc12_subm1, defaultMethod = c("norm", "logreg",
    "polyreg"), seed = 2016, n.imp = 5)
Number of multiple imputations:  5
Missing cells per column:
    seqn riagendr ridreth1 wtmec2yr  sdmvpsu sdmvstra indfmpir   bmxbmi   age18p      age   marcat
```

```
       0        0        0        0        0        0      487       90        0        0      300
 descode bpxdi1_1     agec   agecsq  age4cat
       0      503        0        0        0
Imputation methods:
     seqn  riagendr  ridreth1  wtmec2yr   sdmvpsu  sdmvstra  indfmpir    bmxbmi    age18p       age
       ""        ""        ""        ""        ""        ""    "norm"    "norm"        ""        ""
   marcat   descode  bpxdi1_1      agec    agecsq   age4cat
"polyreg"        ""    "norm"        ""        ""        ""
VisitSequence:
indfmpir   bmxbmi   marcat bpxdi1_1
       7        8       11       13
PredictorMatrix:
         seqn riagendr ridreth1 wtmec2yr sdmvpsu sdmvstra indfmpir bmxbmi age18p age marcat descode
seqn        0        0        0        0       0        0        0      0       0   0      0        0
riagendr    0        0        0        0       0        0        0      0       0   0      0        0
ridreth1    0        0        0        0       0        0        0      0       0   0      0        0
wtmec2yr    0        0        0        0       0        0        0      0       0   0      0        0
sdmvpsu     0        0        0        0       0        0        0      0       0   0      0        0
sdmvstra    0        0        0        0       0        0        0      0       0   0      0        0
indfmpir    1        1        1        1       1        1        0      1       0   1      1        0
bmxbmi      1        1        1        1       1        1        1      0       0   1      1        0
age18p      0        0        0        0       0        0        0      0       0   0      0        0
age         0        0        0        0       0        0        0      0       0   0      0        0
marcat      1        1        1        1       1        1        1      1       0   1      0        0
descode     0        0        0        0       0        0        0      0       0   0      0        0
bpxdi1_1    1        1        1        1       1        1        1      1       0   1      1        0
agec        0        0        0        0       0        0        0      0       0   0      0        0
agecsq      0        0        0        0       0        0        0      0       0   0      0        0
age4cat     0        0        0        0       0        0        0      0       0   0      0        0
         bpxdi1_1 agec agecsq age4cat
seqn            0    0      0       0
riagendr        0    0      0       0
ridreth1        0    0      0       0
wtmec2yr        0    0      0       0
sdmvpsu         0    0      0       0
sdmvstra        0    0      0       0
indfmpir        1    0      1       1
bmxbmi          1    0      1       1
age18p          0    0      0       0
age             0    0      0       0
marcat          1    0      1       1
descode         0    0      0       0
bpxdi1_1        0    0      1       1
agec            0    0      0       0
agecsq          0    0      0       0
age4cat         0    0      0       0
Random generator seed value:  2016
```

```
> # add a predictor matrix to control imputation model predictors for each imputed variable
> pred <- impm1$predictorMatrix
> pred[,"sdmvpsu"] <- 0
> pred[,"sdmvstra"] <- 0
> pred[,"seqn"] <- 0
> pred[,"age18p"] <- 0
> pred[,"descode"] <- 1
> impm1$predictorMatrix
         seqn riagendr ridreth1 wtmec2yr sdmvpsu sdmvstra indfmpir bmxbmi age18p age marcat descode
seqn        0        0        0        0       0        0        0      0      0   0      0        0
riagendr    0        0        0        0       0        0        0      0      0   0      0        0
ridreth1    0        0        0        0       0        0        0      0      0   0      0        0
wtmec2yr    0        0        0        0       0        0        0      0      0   0      0        0
sdmvpsu     0        0        0        0       0        0        0      0      0   0      0        0
sdmvstra    0        0        0        0       0        0        0      0      0   0      0        0
indfmpir    1        1        1        1       1        1        0      1      0   1      1        0
bmxbmi      1        1        1        1       1        1        1      0      0   1      1        0
age18p      0        0        0        0       0        0        0      0      0   0      0        0
age         0        0        0        0       0        0        0      0      0   0      0        0
marcat      1        1        1        1       1        1        1      1      0   1      0        0
descode     0        0        0        0       0        0        0      0      0   0      0        0
bpxdi1_1    1        1        1        1       1        1        1      1      0   1      1        0
agec        0        0        0        0       0        0        0      0      0   0      0        0
agecsq      0        0        0        0       0        0        0      0      0   0      0        0
age4cat     0        0        0        0       0        0        0      0      0   0      0        0
         bpxdi1_1 agec agecsq age4cat
seqn            0    0      0       0
riagendr        0    0      0       0
ridreth1        0    0      0       0
wtmec2yr        0    0      0       0
sdmvpsu         0    0      0       0
sdmvstra        0    0      0       0
indfmpir        1    0      1       1
bmxbmi          1    0      1       1
age18p          0    0      0       0
age             0    0      0       0
marcat          1    0      1       1
descode         0    0      0       0
bpxdi1_1        0    0      1       1
agec            0    0      0       0
agecsq          0    0      0       0
age4cat         0    0      0       0
> pred
         seqn riagendr ridreth1 wtmec2yr sdmvpsu sdmvstra indfmpir bmxbmi age18p age marcat descode
seqn        0        0        0        0       0        0        0      0      0   0      0        1
riagendr    0        0        0        0       0        0        0      0      0   0      0        1
ridreth1    0        0        0        0       0        0        0      0      0   0      0        1
wtmec2yr    0        0        0        0       0        0        0      0      0   0      0        1
sdmvpsu     0        0        0        0       0        0        0      0      0   0      0        1
sdmvstra    0        0        0        0       0        0        0      0      0   0      0        1
indfmpir    0        1        1        1       0        0        0      1      0   1      1        1
bmxbmi      0        1        1        1       0        0        1      0      0   1      1        1
age18p      0        0        0        0       0        0        0      0      0   0      0        1
```

```
age         0        0        0        0       0        0        0      0      0   0      0       1
marcat      0        1        1        1       0        0        1      1      0   1      0       1
descode     0        0        0        0       0        0        0      0      0   0      0       1
bpxdi1_1    0        1        1        1       0        0        1      1      0   1      1       1
agec        0        0        0        0       0        0        0      0      0   0      0       1
agecsq      0        0        0        0       0        0        0      0      0   0      0       1
age4cat     0        0        0        0       0        0        0      0      0   0      0       1
          bpxdi1_1 agec agecsq age4cat
seqn             0    0      0       0
riagendr         0    0      0       0
ridreth1         0    0      0       0
wtmec2yr         0    0      0       0
sdmvpsu          0    0      0       0
sdmvstra         0    0      0       0
indfmpir         1    0      1       1
bmxbmi           1    0      1       1
age18p           0    0      0       0
age              0    0      0       0
marcat           1    0      1       1
descode          0    0      0       0
bpxdi1_1         0    0      1       1
agec             0    0      0       0
agecsq           0    0      0       0
age4cat          0    0      0       0
```

```
> impm1_pred <- mice(nhanesc12_subm1, pred=pred, n.imp=5, seed=2016, defaultMethod=c("norm","logreg","polyreg"))

 iter imp variable
  1   1  indfmpir  bmxbmi  marcat  bpxdi1_1
  1   2  indfmpir  bmxbmi  marcat  bpxdi1_1
  1   3  indfmpir  bmxbmi  marcat  bpxdi1_1
  1   4  indfmpir  bmxbmi  marcat  bpxdi1_1
  1   5  indfmpir  bmxbmi  marcat  bpxdi1_1
  2   1  indfmpir  bmxbmi  marcat  bpxdi1_1
  2   2  indfmpir  bmxbmi  marcat  bpxdi1_1
  2   3  indfmpir  bmxbmi  marcat  bpxdi1_1
  2   4  indfmpir  bmxbmi  marcat  bpxdi1_1
  2   5  indfmpir  bmxbmi  marcat  bpxdi1_1
  3   1  indfmpir  bmxbmi  marcat  bpxdi1_1
  3   2  indfmpir  bmxbmi  marcat  bpxdi1_1
  3   3  indfmpir  bmxbmi  marcat  bpxdi1_1
  3   4  indfmpir  bmxbmi  marcat  bpxdi1_1
  3   5  indfmpir  bmxbmi  marcat  bpxdi1_1
  4   1  indfmpir  bmxbmi  marcat  bpxdi1_1
  4   2  indfmpir  bmxbmi  marcat  bpxdi1_1
  4   3  indfmpir  bmxbmi  marcat  bpxdi1_1
  4   4  indfmpir  bmxbmi  marcat  bpxdi1_1
  4   5  indfmpir  bmxbmi  marcat  bpxdi1_1
  5   1  indfmpir  bmxbmi  marcat  bpxdi1_1
  5   2  indfmpir  bmxbmi  marcat  bpxdi1_1
  5   3  indfmpir  bmxbmi  marcat  bpxdi1_1
```

```
  5   4  indfmpir  bmxbmi  marcat  bpxdi1_1
  5   5  indfmpir  bmxbmi  marcat  bpxdi1_1
> summary(impm1_pred)
Multiply imputed data set
Call:
mice(data = nhanesc12_subm1, predictorMatrix = pred, defaultMethod = c("norm",
    "logreg", "polyreg"), seed = 2016, n.imp = 5)
Number of multiple imputations:  5
Missing cells per column:
    seqn riagendr ridreth1 wtmec2yr  sdmvpsu sdmvstra indfmpir   bmxbmi   age18p      age   marcat
       0        0        0        0        0        0      487       90        0        0      300
 descode bpxdi1_1     agec   agecsq  age4cat
       0      503        0        0        0
Imputation methods:
     seqn  riagendr  ridreth1  wtmec2yr   sdmvpsu  sdmvstra  indfmpir    bmxbmi    age18p       age
       ""        ""        ""        ""        ""        ""    "norm"    "norm"        ""        ""
   marcat   descode  bpxdi1_1      agec    agecsq   age4cat
"polyreg"        ""    "norm"        ""        ""        ""
VisitSequence:
indfmpir   bmxbmi   marcat bpxdi1_1
       7        8       11       13
PredictorMatrix:
         seqn riagendr ridreth1 wtmec2yr sdmvpsu sdmvstra indfmpir bmxbmi age18p age marcat descode
seqn        0        0        0        0       0        0        0      0      0   0      0       0
riagendr    0        0        0        0       0        0        0      0      0   0      0       0
ridreth1    0        0        0        0       0        0        0      0      0   0      0       0
wtmec2yr    0        0        0        0       0        0        0      0      0   0      0       0
sdmvpsu     0        0        0        0       0        0        0      0      0   0      0       0
sdmvstra    0        0        0        0       0        0        0      0      0   0      0       0
indfmpir    0        1        1        1       0        0        0      1      0   1      1       1
bmxbmi      0        1        1        1       0        0        1      0      0   1      1       1
age18p      0        0        0        0       0        0        0      0      0   0      0       0
age         0        0        0        0       0        0        0      0      0   0      0       0
marcat      0        1        1        1       0        0        1      1      0   1      0       1
descode     0        0        0        0       0        0        0      0      0   0      0       0
bpxdi1_1    0        1        1        1       0        0        1      1      0   1      1       1
agec        0        0        0        0       0        0        0      0      0   0      0       0
agecsq      0        0        0        0       0        0        0      0      0   0      0       0
age4cat     0        0        0        0       0        0        0      0      0   0      0       0
         bpxdi1_1 agec agecsq age4cat
seqn            0    0      0       0
riagendr        0    0      0       0
ridreth1        0    0      0       0
wtmec2yr        0    0      0       0
sdmvpsu         0    0      0       0
sdmvstra        0    0      0       0
indfmpir        1    0      1       1
bmxbmi          1    0      1       1
age18p          0    0      0       0
age             0    0      0       0
marcat          1    0      1       1
descode         0    0      0       0
bpxdi1_1        0    0      1       1
```

```
agec            0    0      0       0
agecsq          0    0      0       0
age4cat         0    0      0       0
Random generator seed value:  2016

> # create new high diastolic blood pressure base on imputed bpxdi1_1, prepare long data set and then back to
mids after computation
> longm1 <- complete(impm1_pred, action='long', include=TRUE)
> longm1$high_diastolic <- ifelse(longm1$bpxdi1_1 >=90,1,0)
> summary(longm1)
 .imp          .id             seqn          riagendr        ridreth1        wtmec2yr         sdmvpsu
 0:5615   1      :    6   Min.   :62161   Min.   :1.000   Min.   :1.000   Min.   :  4413   Min.   :1.000
 1:5615   1000   :    6   1st Qu.:64610   1st Qu.:1.000   1st Qu.:3.000   1st Qu.: 16169   1st Qu.:1.000
 2:5615   1001   :    6   Median :67109   Median :2.000   Median :3.000   Median : 24567   Median :2.000
 3:5615   1002   :    6   Mean   :67076   Mean   :1.506   Mean   :3.301   Mean   : 41318   Mean   :1.638
 4:5615   1003   :    6   3rd Qu.:69533   3rd Qu.:2.000   3rd Qu.:4.000   3rd Qu.: 45331   3rd Qu.:2.000
 5:5615   1004   :    6   Max.   :71915   Max.   :2.000   Max.   :5.000   Max.   :222580   Max.   :3.000
          (Other):33654
    sdmvstra         indfmpir          bmxbmi           age18p       age         marcat
 Min.   : 90.00   Min.   :-3.691   Min.   : 8.257   Min.   :1   Min.   :18.00   1   :18312
 1st Qu.: 92.00   1st Qu.: 0.960   1st Qu.:23.800   1st Qu.:1   1st Qu.:31.00   2   : 7134
 Median : 96.00   Median : 1.883   Median :27.400   Median :1   Median :47.00   3   : 7944
 Mean   : 95.87   Mean   : 2.367   Mean   :28.625   Mean   :1   Mean   :47.17   NA's:  300
 3rd Qu.: 99.00   3rd Qu.: 3.890   3rd Qu.:32.100   3rd Qu.:1   3rd Qu.:62.00
 Max.   :103.00   Max.   : 8.608   Max.   :82.100   Max.   :1   Max.   :80.00
                  NA's   :487      NA's   :90
    descode          bpxdi1_1           agec              agecsq             age4cat      high_diastolic
 Min.   : 901.0   Min.   : 10.00   Min.   :-28.3552   Min.   :   0.1261   Min.   :1.000   Min.   :0.0000
 1st Qu.: 922.0   1st Qu.: 64.00   1st Qu.:-15.3552   1st Qu.:  58.4436   1st Qu.:2.000   1st Qu.:0.0000
 Median : 961.0   Median : 72.00   Median :  0.6448   Median : 244.7610   Median :3.000   Median :0.0000
 Mean   : 960.3   Mean   : 70.97   Mean   :  0.8099   Mean   : 345.0635   Mean   :2.594   Mean   :0.0612
 3rd Qu.: 992.0   3rd Qu.: 78.00   3rd Qu.: 15.6448   3rd Qu.: 559.0785   3rd Qu.:3.000   3rd Qu.:0.0000
 Max.   :1032.0   Max.   :120.00   Max.   : 33.6448   Max.   :1131.9752   Max.   :4.000   Max.   :1.0000
                  NA's   :503                                                            NA's   :503

> #use as.mids() to convert back to mids object
> impm1a  <- as.mids(longm1)
>
> # convert mids to data useable for work in mitools
> library(mitools)
> mydatam1 <- imputationList(lapply(1:5, complete, x=impm1a))
> summary(mydatam1)
            Length Class  Mode
imputations 5      -none- list
call        2      -none- call

> # set survey design
> library(survey)
> desm1 <- svydesign(id=~sdmvstra, strat=~sdmvpsu, weight=~wtmec2yr, data=(mydatam1), nest=TRUE)
> summary(desm1)
        Length Class  Mode
designs 5      -none- list
call    6      -none- call
```

```
> # run design based logistic model with svyglm using 5 imputed data sets contained in desm1, Taable 12.5
> fitm1 <- with(desm1, svyglm (high_diastolic ~ factor(ridreth1) + factor(riagendr) + agec + agecsq,
family=quasibinomial))
> summary(MIcombine(fitm1))
Multiple imputation results:
      with(desm1, svyglm(high_diastolic ~ factor(ridreth1) + factor(riagendr) +
    agec + agecsq, family = quasibinomial))
      MIcombine.default(fitm1)
                     results           se      (lower       upper) missInfo
(Intercept)       -2.21225529 0.2363615980 -2.67555595 -1.748954637      2 %
factor(ridreth1)2 -0.65726291 0.2721286781 -1.19131231 -0.123213501      7 %
factor(ridreth1)3  0.14335568 0.2227878149 -0.29346488  0.580176235      4 %
factor(ridreth1)4  0.61474724 0.2373021838  0.14958714  1.079907345      2 %
factor(ridreth1)5  0.08433028 0.2932258468 -0.49178742  0.660447983      9 %
factor(riagendr)2 -0.55414764 0.1769625808 -0.90219639 -0.206098883     11 %
agec               0.01025488 0.0069408411 -0.00335723  0.023866981      5 %
agecsq            -0.00177005 0.0002429929 -0.00224678 -0.001293319      6 %
>
> # combined mean high blood pressure with design adjustment, Table 12.4
> fitm1_mean <- with(desm1, svymean(~factor(high_diastolic), se=T, na.rm=T, ci=T ))
> # mean high blood pressure for each imputed data set
> fitm1_mean
[[1]]
                            mean     SE
factor(high_diastolic)0 0.939158 0.0065
factor(high_diastolic)1 0.060842 0.0065

[[2]]
                            mean     SE
factor(high_diastolic)0 0.937171 0.0061
factor(high_diastolic)1 0.062829 0.0061

[[3]]
                            mean     SE
factor(high_diastolic)0 0.937128 0.0072
factor(high_diastolic)1 0.062872 0.0072

[[4]]
                           mean     SE
factor(high_diastolic)0 0.94144 0.0065
factor(high_diastolic)1 0.05856 0.0065

[[5]]
                            mean     SE
factor(high_diastolic)0 0.939831 0.0064
factor(high_diastolic)1 0.060169 0.0064

attr(,"call")
with(desm1, svymean(~factor(high_diastolic), se = T, na.rm = T,
    ci = T))
> # Use MIcombine for overall combined and design-adjusted mean/se
> summary(MIcombine(fitm1_mean))
Multiple imputation results:
```

```
      with(desm1, svymean(~factor(high_diastolic), se = T, na.rm = T,
    ci = T))
      MIcombine.default(fitm1_mean)
                           results          se     (lower    upper) missInfo
factor(high_diastolic)0 0.93894542 0.006851186 0.92548700 0.9524038      9 %
factor(high_diastolic)1 0.06105458 0.006851186 0.04759615 0.0745130      9 %
>
> # Obtain means for 3 continuous variables imputed using method 1 as well, Table 12.3
> fitm1_ex12_3 <- with(desm1, svymean(~bmxbmi+bpxdi1_1+indfmpir), se=T, na.rm=T, ci=T )
> fitm1_ex12_3
[[1]]
           mean     SE
bmxbmi   28.634 0.2356
bpxdi1_1 71.518 0.4253
indfmpir  2.855 0.1118

[[2]]
            mean     SE
bmxbmi   28.6275 0.2262
bpxdi1_1 71.6151 0.4422
indfmpir  2.8398 0.1126

[[3]]
            mean     SE
bmxbmi   28.6419 0.2314
bpxdi1_1 71.5659 0.4521
indfmpir  2.8425 0.1110

[[4]]
           mean     SE
bmxbmi   28.641 0.2315
bpxdi1_1 71.460 0.4305
indfmpir  2.852 0.1103

[[5]]
            mean     SE
bmxbmi   28.6236 0.2296
bpxdi1_1 71.5747 0.4392
indfmpir  2.8507 0.1122
attr(,"call")
with(desm1, svymean(~bmxbmi + bpxdi1_1 + indfmpir), se = T, na.rm = T,
    ci = T)
```

```
> # Method 2 WITHOUT Design Variables in Model

> # return to original data without the CC high diastolic blood pressure
> nhanesc12_subm2 <- nhanesc12[ which(nhanesc12$age18p==1 & nhanesc12$wtmec2yr >0),]

> #nhanesc12_subm2$descode=factor(nhanesc12_subm2$descode)
> summary(nhanesc12_subm2)
      seqn          riagendr        ridreth1        wtmec2yr         sdmvpsu         sdmvstra
 Min.   :62161   Min.   :1.000   Min.   :1.000   Min.   :  4413   Min.   :1.000   Min.   : 90.00
 1st Qu.:64611   1st Qu.:1.000   1st Qu.:3.000   1st Qu.: 16174   1st Qu.:1.000   1st Qu.: 92.00
 Median :67109   Median :2.000   Median :3.000   Median : 24567   Median :2.000   Median : 96.00
 Mean   :67076   Mean   :1.506   Mean   :3.301   Mean   : 41318   Mean   :1.638   Mean   : 95.87
 3rd Qu.:69533   3rd Qu.:2.000   3rd Qu.:4.000   3rd Qu.: 45238   3rd Qu.:2.000   3rd Qu.: 99.00
 Max.   :71915   Max.   :2.000   Max.   :5.000   Max.   :222580   Max.   :3.000   Max.   :103.00

    indfmpir         bmxbmi          age18p       age         marcat        descode          bpxdi1_1
 Min.   :0.000   Min.   :13.40   Min.   :1   Min.   :18.00   1   :2991   Min.   : 901.0   Min.   : 10.00
 1st Qu.:0.950   1st Qu.:23.80   1st Qu.:1   1st Qu.:31.00   2   :1183   1st Qu.: 922.0   1st Qu.: 64.00
 Median :1.840   Median :27.40   Median :1   Median :47.00   3   :1141   Median : 961.0   Median : 72.00
 Mean   :2.374   Mean   :28.62   Mean   :1   Mean   :47.17   NA's: 300   Mean   : 960.3   Mean   : 71.02
 3rd Qu.:3.933   3rd Qu.:32.00   3rd Qu.:1   3rd Qu.:62.00               3rd Qu.: 992.0   3rd Qu.: 78.00
 Max.   :5.000   Max.   :82.10   Max.   :1   Max.   :80.00               Max.   :1032.0   Max.   :120.00
 NA's   :487     NA's   :90                                                             NA's   :503
      agec             agecsq            age4cat
 Min.   :-28.3552   Min.   :   0.1261   Min.   :1.000
 1st Qu.:-15.3552   1st Qu.:  58.4436   1st Qu.:2.000
 Median :  0.6448   Median : 244.7610   Median :3.000
 Mean   :  0.8099   Mean   : 345.0635   Mean   :2.594
 3rd Qu.: 15.6448   3rd Qu.: 559.0785   3rd Qu.:3.000
 Max.   : 33.6448   Max.   :1131.9752   Max.   :4.000


> # use mice to impute and specify type of default method for imputation models
> # run without custom predictor matrix first
> impm2 <- mice(nhanesc12_subm2, n.imp=5, seed=2016, defaultMethod=c("norm","logreg","polyreg"))

 iter imp variable
  1   1  indfmpir  bmxbmi  marcat  bpxdi1_1
  1   2  indfmpir  bmxbmi  marcat  bpxdi1_1
  1   3  indfmpir  bmxbmi  marcat  bpxdi1_1
  1   4  indfmpir  bmxbmi  marcat  bpxdi1_1
  1   5  indfmpir  bmxbmi  marcat  bpxdi1_1
  2   1  indfmpir  bmxbmi  marcat  bpxdi1_1
  2   2  indfmpir  bmxbmi  marcat  bpxdi1_1
  2   3  indfmpir  bmxbmi  marcat  bpxdi1_1
  2   4  indfmpir  bmxbmi  marcat  bpxdi1_1
  2   5  indfmpir  bmxbmi  marcat  bpxdi1_1
  3   1  indfmpir  bmxbmi  marcat  bpxdi1_1
  3   2  indfmpir  bmxbmi  marcat  bpxdi1_1
  3   3  indfmpir  bmxbmi  marcat  bpxdi1_1
  3   4  indfmpir  bmxbmi  marcat  bpxdi1_1
  3   5  indfmpir  bmxbmi  marcat  bpxdi1_1
  4   1  indfmpir  bmxbmi  marcat  bpxdi1_1
```

```
  4   2  indfmpir  bmxbmi  marcat  bpxdi1_1
  4   3  indfmpir  bmxbmi  marcat  bpxdi1_1
  4   4  indfmpir  bmxbmi  marcat  bpxdi1_1
  4   5  indfmpir  bmxbmi  marcat  bpxdi1_1
  5   1  indfmpir  bmxbmi  marcat  bpxdi1_1
  5   2  indfmpir  bmxbmi  marcat  bpxdi1_1
  5   3  indfmpir  bmxbmi  marcat  bpxdi1_1
  5   4  indfmpir  bmxbmi  marcat  bpxdi1_1
  5   5  indfmpir  bmxbmi  marcat  bpxdi1_1
> impm2$predictorMatrix
         seqn riagendr ridreth1 wtmec2yr sdmvpsu sdmvstra indfmpir bmxbmi age18p age marcat descode
seqn        0        0        0        0       0        0        0      0      0   0      0        0
riagendr    0        0        0        0       0        0        0      0      0   0      0        0
ridreth1    0        0        0        0       0        0        0      0      0   0      0        0
wtmec2yr    0        0        0        0       0        0        0      0      0   0      0        0
sdmvpsu     0        0        0        0       0        0        0      0      0   0      0        0
sdmvstra    0        0        0        0       0        0        0      0      0   0      0        0
indfmpir    1        1        1        1       1        1        0      1      0   1      1        0
bmxbmi      1        1        1        1       1        1        1      0      0   1      1        0
age18p      0        0        0        0       0        0        0      0      0   0      0        0
age         0        0        0        0       0        0        0      0      0   0      0        0
marcat      1        1        1        1       1        1        1      1      0   1      0        0
descode     0        0        0        0       0        0        0      0      0   0      0        0
bpxdi1_1    1        1        1        1       1        1        1      1      0   1      1        0
agec        0        0        0        0       0        0        0      0      0   0      0        0
agecsq      0        0        0        0       0        0        0      0      0   0      0        0
age4cat     0        0        0        0       0        0        0      0      0   0      0        0
         bpxdi1_1 agec agecsq age4cat
seqn            0    0      0       0
riagendr        0    0      0       0
ridreth1        0    0      0       0
wtmec2yr        0    0      0       0
sdmvpsu         0    0      0       0
sdmvstra        0    0      0       0
indfmpir        1    0      1       1
bmxbmi          1    0      1       1
age18p          0    0      0       0
age             0    0      0       0
marcat          1    0      1       1
descode         0    0      0       0
bpxdi1_1        0    0      1       1
agec            0    0      0       0
agecsq          0    0      0       0
age4cat         0    0      0       0
> summary(impm2)
Multiply imputed data set
Call:
mice(data = nhanesc12_subm2, defaultMethod = c("norm", "logreg",
    "polyreg"), seed = 2016, n.imp = 5)
Number of multiple imputations:  5
Missing cells per column:
    seqn riagendr ridreth1 wtmec2yr  sdmvpsu sdmvstra indfmpir   bmxbmi   age18p      age   marcat
       0        0        0        0        0        0      487       90        0        0      300
```

```
 descode bpxdi1_1     agec   agecsq  age4cat
       0      503        0        0        0
Imputation methods:
     seqn  riagendr  ridreth1  wtmec2yr   sdmvpsu  sdmvstra  indfmpir    bmxbmi    age18p       age
       ""        ""        ""        ""        ""        ""    "norm"    "norm"        ""        ""
   marcat   descode  bpxdi1_1      agec    agecsq   age4cat
"polyreg"        ""    "norm"        ""        ""        ""
VisitSequence:
indfmpir   bmxbmi   marcat bpxdi1_1
       7        8       11       13
PredictorMatrix:
         seqn riagendr ridreth1 wtmec2yr sdmvpsu sdmvstra indfmpir bmxbmi age18p age marcat descode
seqn        0        0        0        0       0        0        0      0      0   0      0        0
riagendr    0        0        0        0       0        0        0      0      0   0      0        0
ridreth1    0        0        0        0       0        0        0      0      0   0      0        0
wtmec2yr    0        0        0        0       0        0        0      0      0   0      0        0
sdmvpsu     0        0        0        0       0        0        0      0      0   0      0        0
sdmvstra    0        0        0        0       0        0        0      0      0   0      0        0
indfmpir    1        1        1        1       1        1        0      1      0   1      1        0
bmxbmi      1        1        1        1       1        1        1      0      0   1      1        0
age18p      0        0        0        0       0        0        0      0      0   0      0        0
age         0        0        0        0       0        0        0      0      0   0      0        0
marcat      1        1        1        1       1        1        1      1      0   1      0        0
descode     0        0        0        0       0        0        0      0      0   0      0        0
bpxdi1_1    1        1        1        1       1        1        1      1      0   1      1        0
agec        0        0        0        0       0        0        0      0      0   0      0        0
agecsq      0        0        0        0       0        0        0      0      0   0      0        0
age4cat     0        0        0        0       0        0        0      0      0   0      0        0
         bpxdi1_1 agec agecsq age4cat
seqn            0    0      0       0
riagendr        0    0      0       0
ridreth1        0    0      0       0
wtmec2yr        0    0      0       0
sdmvpsu         0    0      0       0
sdmvstra        0    0      0       0
indfmpir        1    0      1       1
bmxbmi          1    0      1       1
age18p          0    0      0       0
age             0    0      0       0
marcat          1    0      1       1
descode         0    0      0       0
bpxdi1_1        0    0      1       1
agec            0    0      0       0
agecsq          0    0      0       0
age4cat         0    0      0       0
Random generator seed value:  2016
>
> # add a predictor matrix to control imputation model predictors for each imputed variable
> pred <- impm2$predictorMatrix
> pred[,"sdmvpsu"] <- 0
> pred[,"sdmvstra"] <- 0
> pred[,"seqn"] <- 0
> pred[,"age18p"] <- 0
```

```
> pred[,"descode"] <- 0
> pred[,"wtmec2yr"] <- 0
> impm2$predictorMatrix
         seqn riagendr ridreth1 wtmec2yr sdmvpsu sdmvstra indfmpir bmxbmi age18p age marcat descode
seqn        0        0        0        0       0        0        0      0      0   0      0        0
riagendr    0        0        0        0       0        0        0      0      0   0      0        0
ridreth1    0        0        0        0       0        0        0      0      0   0      0        0
wtmec2yr    0        0        0        0       0        0        0      0      0   0      0        0
sdmvpsu     0        0        0        0       0        0        0      0      0   0      0        0
sdmvstra    0        0        0        0       0        0        0      0      0   0      0        0
indfmpir    1        1        1        1       1        1        0      1      0   1      1        0
bmxbmi      1        1        1        1       1        1        1      0      0   1      1        0
age18p      0        0        0        0       0        0        0      0      0   0      0        0
age         0        0        0        0       0        0        0      0      0   0      0        0
marcat      1        1        1        1       1        1        1      1      0   1      0        0
descode     0        0        0        0       0        0        0      0      0   0      0        0
bpxdi1_1    1        1        1        1       1        1        1      1      0   1      1        0
agec        0        0        0        0       0        0        0      0      0   0      0        0
agecsq      0        0        0        0       0        0        0      0      0   0      0        0
age4cat     0        0        0        0       0        0        0      0      0   0      0        0
         bpxdi1_1 agec agecsq age4cat
seqn            0    0      0       0
riagendr        0    0      0       0
ridreth1        0    0      0       0
wtmec2yr        0    0      0       0
sdmvpsu         0    0      0       0
sdmvstra        0    0      0       0
indfmpir        1    0      1       1
bmxbmi          1    0      1       1
age18p          0    0      0       0
age             0    0      0       0
marcat          1    0      1       1
descode         0    0      0       0
bpxdi1_1        0    0      1       1
agec            0    0      0       0
agecsq          0    0      0       0
age4cat         0    0      0       0
> pred
         seqn riagendr ridreth1 wtmec2yr sdmvpsu sdmvstra indfmpir bmxbmi age18p age marcat descode
seqn        0        0        0        0       0        0        0      0      0   0      0        0
riagendr    0        0        0        0       0        0        0      0      0   0      0        0
ridreth1    0        0        0        0       0        0        0      0      0   0      0        0
wtmec2yr    0        0        0        0       0        0        0      0      0   0      0        0
sdmvpsu     0        0        0        0       0        0        0      0      0   0      0        0
sdmvstra    0        0        0        0       0        0        0      0      0   0      0        0
indfmpir    0        1        1        0       0        0        0      1      0   1      1        0
bmxbmi      0        1        1        0       0        0        1      0      0   1      1        0
age18p      0        0        0        0       0        0        0      0      0   0      0        0
age         0        0        0        0       0        0        0      0      0   0      0        0
marcat      0        1        1        0       0        0        1      1      0   1      0        0
descode     0        0        0        0       0        0        0      0      0   0      0        0
bpxdi1_1    0        1        1        0       0        0        1      1      0   1      1        0
agec        0        0        0        0       0        0        0      0      0   0      0        0
```

```
agecsq      0       0       0       0       0       0       0      0      0   0      0       0
age4cat     0       0       0       0       0       0       0      0      0   0      0       0
         bpxdi1_1 agec agecsq age4cat
seqn            0    0      0       0
riagendr        0    0      0       0
ridreth1        0    0      0       0
wtmec2yr        0    0      0       0
sdmvpsu         0    0      0       0
sdmvstra        0    0      0       0
indfmpir        1    0      1       1
bmxbmi          1    0      1       1
age18p          0    0      0       0
age             0    0      0       0
marcat          1    0      1       1
descode         0    0      0       0
bpxdi1_1        0    0      1       1
agec            0    0      0       0
agecsq          0    0      0       0
age4cat         0    0      0       0

> impm2_pred <- mice(nhanesc12_subm2, pred=pred, n.imp=5, seed=2016, defaultMethod=c("norm","logreg","polyreg"))

 iter imp variable
  1   1  indfmpir  bmxbmi  marcat  bpxdi1_1
  1   2  indfmpir  bmxbmi  marcat  bpxdi1_1
  1   3  indfmpir  bmxbmi  marcat  bpxdi1_1
  1   4  indfmpir  bmxbmi  marcat  bpxdi1_1
  1   5  indfmpir  bmxbmi  marcat  bpxdi1_1
  2   1  indfmpir  bmxbmi  marcat  bpxdi1_1
  2   2  indfmpir  bmxbmi  marcat  bpxdi1_1
  2   3  indfmpir  bmxbmi  marcat  bpxdi1_1
  2   4  indfmpir  bmxbmi  marcat  bpxdi1_1
  2   5  indfmpir  bmxbmi  marcat  bpxdi1_1
  3   1  indfmpir  bmxbmi  marcat  bpxdi1_1
  3   2  indfmpir  bmxbmi  marcat  bpxdi1_1
  3   3  indfmpir  bmxbmi  marcat  bpxdi1_1
  3   4  indfmpir  bmxbmi  marcat  bpxdi1_1
  3   5  indfmpir  bmxbmi  marcat  bpxdi1_1
  4   1  indfmpir  bmxbmi  marcat  bpxdi1_1
  4   2  indfmpir  bmxbmi  marcat  bpxdi1_1
  4   3  indfmpir  bmxbmi  marcat  bpxdi1_1
  4   4  indfmpir  bmxbmi  marcat  bpxdi1_1
  4   5  indfmpir  bmxbmi  marcat  bpxdi1_1
  5   1  indfmpir  bmxbmi  marcat  bpxdi1_1
  5   2  indfmpir  bmxbmi  marcat  bpxdi1_1
  5   3  indfmpir  bmxbmi  marcat  bpxdi1_1
  5   4  indfmpir  bmxbmi  marcat  bpxdi1_1
  5   5  indfmpir  bmxbmi  marcat  bpxdi1_1
```

```
> summary(impm2_pred)
Multiply imputed data set
Call:
mice(data = nhanesc12_subm2, predictorMatrix = pred, defaultMethod = c("norm",
    "logreg", "polyreg"), seed = 2016, n.imp = 5)
Number of multiple imputations:  5
Missing cells per column:
    seqn riagendr ridreth1 wtmec2yr  sdmvpsu sdmvstra indfmpir   bmxbmi   age18p      age   marcat
       0        0        0        0        0        0      487       90        0        0      300
 descode bpxdi1_1     agec   agecsq  age4cat
       0      503        0        0        0
Imputation methods:
     seqn  riagendr  ridreth1  wtmec2yr   sdmvpsu  sdmvstra  indfmpir    bmxbmi    age18p       age
       ""        ""        ""        ""        ""        ""    "norm"    "norm"        ""        ""
   marcat   descode  bpxdi1_1      agec    agecsq   age4cat
"polyreg"        ""    "norm"        ""        ""        ""
VisitSequence:
indfmpir   bmxbmi   marcat bpxdi1_1
       7        8       11       13
PredictorMatrix:
         seqn riagendr ridreth1 wtmec2yr sdmvpsu sdmvstra indfmpir bmxbmi age18p age marcat descode
seqn        0        0        0        0       0        0        0      0       0   0      0       0
riagendr    0        0        0        0       0        0        0      0       0   0      0       0
ridreth1    0        0        0        0       0        0        0      0       0   0      0       0
wtmec2yr    0        0        0        0       0        0        0      0       0   0      0       0
sdmvpsu     0        0        0        0       0        0        0      0       0   0      0       0
sdmvstra    0        0        0        0       0        0        0      0       0   0      0       0
indfmpir    0        1        1        0       0        0        0      1       0   1      1       0
bmxbmi      0        1        1        0       0        0        1      0       0   1      1       0
age18p      0        0        0        0       0        0        0      0       0   0      0       0
age         0        0        0        0       0        0        0      0       0   0      0       0
marcat      0        1        1        0       0        0        1      1       0   1      0       0
descode     0        0        0        0       0        0        0      0       0   0      0       0
bpxdi1_1    0        1        1        0       0        0        1      1       0   1      1       0
agec        0        0        0        0       0        0        0      0       0   0      0       0
agecsq      0        0        0        0       0        0        0      0       0   0      0       0
age4cat     0        0        0        0       0        0        0      0       0   0      0       0
         bpxdi1_1 agec agecsq age4cat
seqn            0    0      0       0
riagendr        0    0      0       0
ridreth1        0    0      0       0
wtmec2yr        0    0      0       0
sdmvpsu         0    0      0       0
sdmvstra        0    0      0       0
indfmpir        1    0      1       1
bmxbmi          1    0      1       1
age18p          0    0      0       0
age             0    0      0       0
marcat          1    0      1       1
descode         0    0      0       0
bpxdi1_1        0    0      1       1
agec            0    0      0       0
agecsq          0    0      0       0
```

```
age4cat         0    0      0       0
Random generator seed value:  2016
>
> # create new high diastolic blood pressure base on imputed bpxdi1_1, prepare long data set and then back to
mids after computation
> longm2 <- complete(impm2_pred, action='long', include=TRUE)
> longm2$high_diastolic <- ifelse(longm2$bpxdi1_1 >=90,1,0)
> summary(longm2)
 .imp          .id             seqn          riagendr        ridreth1        wtmec2yr         sdmvpsu
 0:5615   1      :    6   Min.   :62161   Min.   :1.000   Min.   :1.000   Min.   :  4413   Min.   :1.000
 1:5615   1000   :    6   1st Qu.:64610   1st Qu.:1.000   1st Qu.:3.000   1st Qu.: 16169   1st Qu.:1.000
 2:5615   1001   :    6   Median :67109   Median :2.000   Median :3.000   Median : 24567   Median :2.000
 3:5615   1002   :    6   Mean   :67076   Mean   :1.506   Mean   :3.301   Mean   : 41318   Mean   :1.638
 4:5615   1003   :    6   3rd Qu.:69533   3rd Qu.:2.000   3rd Qu.:4.000   3rd Qu.: 45331   3rd Qu.:2.000
 5:5615   1004   :    6   Max.   :71915   Max.   :2.000   Max.   :5.000   Max.   :222580   Max.   :3.000
          (Other):33654
    sdmvstra         indfmpir          bmxbmi           age18p       age         marcat
 Min.   : 90.00   Min.   :-3.839   Min.   : 4.303   Min.   :1   Min.   :18.00   1   :18277
 1st Qu.: 92.00   1st Qu.: 0.960   1st Qu.:23.800   1st Qu.:1   1st Qu.:31.00   2   : 7129
 Median : 96.00   Median : 1.890   Median :27.400   Median :1   Median :47.00   3   : 7984
 Mean   : 95.87   Mean   : 2.378   Mean   :28.615   Mean   :1   Mean   :47.17   NA's:  300
 3rd Qu.: 99.00   3rd Qu.: 3.900   3rd Qu.:32.100   3rd Qu.:1   3rd Qu.:62.00
 Max.   :103.00   Max.   : 9.092   Max.   :82.100   Max.   :1   Max.   :80.00
                  NA's   :487      NA's   :90
    descode          bpxdi1_1           agec              agecsq             age4cat      high_diastolic
 Min.   : 901.0   Min.   : 10.00   Min.   :-28.3552   Min.   :   0.1261   Min.   :1.000   Min.   :0.000
 1st Qu.: 922.0   1st Qu.: 64.00   1st Qu.:-15.3552   1st Qu.:  58.4436   1st Qu.:2.000   1st Qu.:0.000
 Median : 961.0   Median : 72.00   Median :  0.6448   Median : 244.7610   Median :3.000   Median :0.000
 Mean   : 960.3   Mean   : 70.97   Mean   :  0.8099   Mean   : 345.0635   Mean   :2.594   Mean   :0.061
 3rd Qu.: 992.0   3rd Qu.: 78.00   3rd Qu.: 15.6448   3rd Qu.: 559.0785   3rd Qu.:3.000   3rd Qu.:0.000
 Max.   :1032.0   Max.   :120.00   Max.   : 33.6448   Max.   :1131.9752   Max.   :4.000   Max.   :1.000
                  NA's   :503                                                            NA's   :503
>
> #use as.mids() to convert back to mids object
> impm2a  <- as.mids(longm2)
>
> # convert mids to data useable for work in mitools
> library(mitools)
> mydatam2 <- imputationList(lapply(1:5, complete, x=impm2a))
> summary(mydatam2)
            Length Class  Mode
imputations 5      -none- list
call        2      -none- call
>
> # set survey design
> library(survey)
> desm2 <- svydesign(id=~sdmvstra, strat=~sdmvpsu, weight=~wtmec2yr, data=(mydatam2), nest=TRUE)
> summary(desm2)
        Length Class  Mode
designs 5      -none- list
call    6      -none- call
>
> # run design based logistic model with svyglm using 5 imputed data sets contained in desm1, Taable 12.5
```

```
> fitm2 <- with(desm2, svyglm (high_diastolic ~ factor(ridreth1) + factor(riagendr) + agec + agecsq,
family=quasibinomial))
> summary(MIcombine(fitm2))
Multiple imputation results:
      with(desm2, svyglm(high_diastolic ~ factor(ridreth1) + factor(riagendr) +
    agec + agecsq, family = quasibinomial))
      MIcombine.default(fitm2)
                      results           se       (lower       upper) missInfo
(Intercept)       -2.276264452 0.2340423016 -2.735019570 -1.817509333      2 %
factor(ridreth1)2 -0.658567994 0.2620548085 -1.174041445 -0.143094542     11 %
factor(ridreth1)3  0.153752272 0.2246763934 -0.286842571  0.594347114      4 %
factor(ridreth1)4  0.658261229 0.2387780360  0.190255632  1.126266826      1 %
factor(ridreth1)5  0.088040408 0.2988129120 -0.498033034  0.674113851      5 %
factor(riagendr)2 -0.501290427 0.1860284829 -0.868953631 -0.133627222     18 %
agec               0.008409436 0.0068413544 -0.004999598  0.021818470      1 %
agecsq            -0.001708093 0.0002510349 -0.002208132 -0.001208054     25 %
>
> # combined mean high blood pressure with design adjustment, Table 12.4
> fitm2_mean <- with(desm2, svymean(~factor(high_diastolic), se=T, na.rm=T, ci=T ))
> # mean high blood pressure for each imputed data set
> fitm2_mean
[[1]]
                            mean     SE
factor(high_diastolic)0 0.938486 0.0067
factor(high_diastolic)1 0.061514 0.0067

[[2]]
                            mean     SE
factor(high_diastolic)0 0.939995 0.0066
factor(high_diastolic)1 0.060005 0.0066

[[3]]
                           mean     SE
factor(high_diastolic)0 0.94009 0.0067
factor(high_diastolic)1 0.05991 0.0067

[[4]]
                           mean     SE
factor(high_diastolic)0 0.94255 0.0065
factor(high_diastolic)1 0.05745 0.0065

[[5]]
                            mean     SE
factor(high_diastolic)0 0.938556 0.0064
factor(high_diastolic)1 0.061444 0.0064

attr(,"call")
with(desm2, svymean(~factor(high_diastolic), se = T, na.rm = T,
    ci = T))
```

```
> # Use MIcombine for overall combined and design-adjusted mean/se
> summary(MIcombine(fitm2_mean))
Multiple imputation results:
      with(desm2, svymean(~factor(high_diastolic), se = T, na.rm = T,
    ci = T))
      MIcombine.default(fitm2_mean)
                          results          se    (lower     upper) missInfo
factor(high_diastolic)0 0.93993534 0.006816101 0.9265561 0.95331460      7 %
factor(high_diastolic)1 0.06006466 0.006816101 0.0466854 0.07344392      7 %
```

> # Note: FEFI method available in R as of early June 2017, see https://sites.google.com/view/jaekwangkim/software for more information, this will be included on ASDA website in the near future.