

MCBS_within_survey_imputation_readme.txt

Purpose: Within survey imputation for missing values in survey variables contained in derived datasets. Creates 5 multiply imputed datasets by population for use in disease calibration step.

Community population--Program/code required:

1. impute_mcbs09.sas
2. mcbs09.set

Institutionalized population--Program/code required:

1. impute_mcbs_09_inst.sas
2. mcbs09_inst.set
3. mcbs09_inst_2a.set
4. mcbs09_inst_2b.set
5. mcbs09_inst_2c.set
6. mcbs09_inst_2d.set
7. mcbs09_inst_2e.set

IVEWare is required to run these data processing steps.

NHANES_within_survey_imputation_readme.txt

Purpose: Within survey imputation for missing values in survey variables.
Creates 5 multiply imputed datasets for use in disease calibration step.

Program/code required:

1. impute_nhanes_65_0910.sas
2. nhanes_65_0910.set

IVWare is required to run these data processing steps.

```

/*****
Program: 'impute_mcbs09.sas'
Purpose: Impute within survey missing values
Data in: Insert path for input datasets
Data out: Insert path for output dataset
*****/

Libname g 'Insert file path'; *For use in calibrations;
Libname w 'Insert file path'; *For use in calculation of HMO weights;

%include "Insert file path\formats_02_01_11.sas";
    options ; *your options statement;
libname oth "Insert file path"; *location of formats and labels;
%include 'Insert file path\xdelete.sas';
%xdelete(ALL);

Data mcbs_ni;
Set g.MCBS_ni09;
lweight=log(weightkg);
drop institution;
Run;

options set=srclib "Insert path for \srclib" sasautos=('!SRCLIB' sasautos)
mautosource ;
%impute(name=mcbs09, dir=Insert file path,setup=old);
%putdata(name=mcbs09, dir=Insert file path,mult=2, dataout=end2);
    run;
%putdata(name=mcbs09, dir=Insert file path,mult=3, dataout=end3);
    run;
%putdata(name=mcbs09, dir=Insert file path,mult=4, dataout=end4);
    run;
%putdata(name=mcbs09, dir=Insert file path,mult=5, dataout=end5);
    run;

/*All subjects are selected to estimate HMO weights*/;
Data w.mcbs_ni_i09; *For HMO weights;
    set end1 end2 end3 end4 end5;
If male=0 then do;
    cgarsr11=0;
    cgarsr82=0;
    PSA1yr=0;
End;
If male=1 then do;
    hyst=0;
    cgarsr9=0;
    cgarsr10=0;

```

```

        PAP_SMEAR=0;
        MAMMOGRAM=0;
    End;
    If everSmoke=0 then smokenow=0;
If wgt>0;
weightkg=exp(lweight);
drop lweight cgar1-cgar125;
Run;

/*Subsets subjects with "Pure" Medicare for future calibration*/
Data g.mcbs_ni_i09;
set end1 end2 end3 end4 end5;
If pure=1;
If male=0 then do;
    cgarsr11=0;
    cgarsr82=0;
    PSA1yr=0;
End;
If male=1 then do;
    hyst=0;
    cgarsr9=0;
    cgarsr10=0;
    PAP_SMEAR=0;
    MAMMOGRAM=0;
End;
    If everSmoke=0 then smokenow=0;
If wgt>0;
weightkg=exp(lweight);
drop lweight cgar1-cgar125 pure mcab_months_3grp mcadv_months;
Run;

```

```

/*****
Program: 'impute_mcbs09.sas'
Purpose: Impute within survey missing values for institutionalized
population
Data in: Insert path for input datasets
Data out: Insert path for output dataset
*****/
%let Y=09;
%let Path = Insert path;
libname sharedi "Insert file path";
libname oth "Insert file path"; *location of formats and labels;

proc contents data=sharedi.inst_mcbs_09;
run;

proc means data=sharedi.inst_mcbs_09 n noprint;
output out=means09;
run;

data means09;
set means09;
if _stat_ = 'N';
run;

proc transpose data=means09 out=means09trans;
by _type_;
run;

data means09trans;
set means09trans;
missing = (1 - (col1/661))*100;
run;

proc sort data=means09trans;
by descending missing;
run;

proc print data=means09trans ;
var _type_ _name_ col1 missing;
run;

/*read in 09 institutionalized data and subset
- exclude variables not of interest
-merge back at end*/
Data mcbs_inst_09;

```

```

Set sharedi.inst_mcbs_09;

if height < 0 then height = .;

/*drop variables not of interest*/
drop cgar1-cgar105 cgar122-cgar125 strat psu wgt
inpatnights inpatstays faccost instcost i_days
institution nbrpeopl type;
Run;

proc contents data=mcbs_inst_09;
run;

/*****imputation code*****/
/*continous variables(default) - weightkg age height */;
options set=srclib "Insert path for \srclib" sasautos=('!SRCLIB' sasautos)
mautosource;
%impute(name=mcbs09_inst, dir=Insert file path,setup=old);
%putdata(name=mcbs09_inst, dir=Insert file path,mult=2,dataout=end2);
run;
%putdata(name=mcbs09_inst, dir=Insert file path,mult=3,dataout=end3);
run;
%putdata(name=mcbs09_inst, dir=Insert file path,mult=4, dataout=end4);
run;
%putdata(name=mcbs09_inst, dir=Insert file path,mult=5, dataout=end5);
run;

data temp;
set end1 end2 end3 end4 end5;
if male = 0 then do;
psa1yr = 0;
end;
if male = 1 then do;
hyst=0;
pap_smear=0;
mammogram=0;
end;
if smoking = 0 then do;
eversmoke = 0;
smokenow = 0;
end;
if smoking = 1 then do;
eversmoke = 1;
smokenow = 0;
end;
if smoking = 2 then do;

```

```

        ever smoke = 1;
        smokenow = 1;
end;

run;

/*diagnostics
- check means in iveware output
- check frequencies in iveware output
- double check frequencies for gender related and smoking vars
*/
options orientation=landscape;
ods rtf file='Insert file path\imp_freqs_1.rtf' columns=2 ;

proc means data=sharedi.inst_mcbs_09 nolabels;
var height weightkg cost ;
run;

proc means data=temp nolabels;
var height weightkg cost ;
class _mult_;
run;

proc freq data=sharedi.inst_mcbs_09;
table didserv ;
run;

proc freq data=temp;
table didserv*_mult_/ norow nopercnt;
run;

proc freq data=sharedi.inst_mcbs_09;
table marital ;
run;

proc freq data=temp;
table marital*_mult_/ norow nopercnt;
run;

proc freq data=sharedi.inst_mcbs_09;
table ed5/ ;
run;

proc freq data=temp;
table ed5*_mult_/ norow nopercnt;
run;

```

```
proc freq data=sharedi.inst_mcbs_09;  
table pneushot/ ;  
run;
```

```
proc freq data=temp;  
table pneushot*_mult_/ norow nopercnt;  
run;
```

```
proc freq data=sharedi.inst_mcbs_09;  
table comphealth/ ;  
run;
```

```
proc freq data=temp;  
table comphealth*_mult_/ norow nopercnt;  
run;
```

```
proc freq data=sharedi.inst_mcbs_09;  
table healthstat/ ;  
run;
```

```
proc freq data=temp;  
table healthstat*_mult_/norow nopercnt;  
run;
```

```
proc freq data=sharedi.inst_mcbs_09;  
table hearingaid/ ;  
run;
```

```
proc freq data=temp;  
table hearingaid*_mult_/ norow nopercnt;  
run;
```

```
proc freq data=sharedi.inst_mcbs_09;  
table hear_inst/ ;  
run;
```

```
proc freq data=temp;  
table hear_inst*_mult_/ norow nopercnt;  
run;
```

```
proc freq data=sharedi.inst_mcbs_09;  
table smoking/ ;  
run;
```

```
proc freq data=temp;
```



```
table smoking*_mult_/ norow nopercent;
run;

proc freq data=sharedi.inst_mcbs_09;
table dif_lift ;
run;

proc freq data=temp;
table dif_lift*_mult_/ norow nopercent;
run;

proc freq data=sharedi.inst_mcbs_09;
table dif_stoop ;
run;

proc freq data=temp;
table dif_stoop*_mult_/norow nopercent;
run;

proc freq data=sharedi.inst_mcbs_09;
table dif_walk ;
run;

proc freq data=temp;
table dif_walk*_mult_/ norow nopercent;
run;

proc freq data=sharedi.inst_mcbs_09;
table prb_dres/ ;
run;

proc freq data=temp;
table prb_dres*_mult_/ norow nopercent;
run;

proc freq data=sharedi.inst_mcbs_09;
table prb_eat/ ;
run;

proc freq data=temp;
table prb_eat*_mult_/ norow nopercent;
run;

proc freq data=sharedi.inst_mcbs_09;
table male*psa1yr/nopercent nocol ;
run;
```

```

proc freq data=temp;
table male*psa1yr*_mult_/ norow nopercnt;
run;

proc freq data=sharedi.inst_mcbs_09;
table male*pap_smear/ nopercnt nocol;
run;

proc freq data=temp;
table male*pap_smear*_mult_/ norow nopercnt;
run;

proc freq data=sharedi.inst_mcbs_09;
table male*hyst/ nopercnt nocol;
run;

proc freq data=temp;
table male*hyst*_mult_/ norow nopercnt;
run;

proc freq data=sharedi.inst_mcbs_09;
table male*mammogram/ nopercnt nocol;
run;

proc freq data=temp;
table male*mammogram*_mult_/ norow nopercnt;
run;

proc freq data=sharedi.inst_mcbs_09;
table cgarsr43/ ;
run;

proc freq data=temp;
table cgarsr43*_mult_/ norow nopercnt;
run;
ods rtf close;

/*****merge in NI data to impute povcat pneushot
flushot*****/
/*read in non institutionalized data*/
libname shared 'Insert file path';
libname wts 'Insert file path';

/*temporary imputed non instutionalized*/
data noninst (rename=(cataract=cgarsr43));

```

```

    set wts.mcbs_ni_i&Y.;
i = 0;
run;

/*temporary institutionalized from imputation stage 1*/
data temp;
set temp;
i = 1;
run;

/*stack institutionalized and non institutionalized for analysis
add labels*/
data stack;
set temp noninst;
drop cgar1-cgar105 cgar122-cgar125 ;
run;

/*assess variables common to both NI and I samples*/
proc means data=stack nolabels noprint;
output out=meanstack;
run;

data meanstack;
set meanstack;
if _stat_ = 'N';
run;

proc transpose data=meanstack out=meanstacktrans;
by _type_;
run;

data meanstacktrans;
set meanstacktrans;
missing = (1 - (col1/43410))*100;
run;

proc sort data=meanstacktrans;
by descending missing;
run;

proc print data=meanstacktrans;
var _type_ _name_ col1 missing;
run;

/*stack institutionalized and non institutionalized for analysis

```

```

add labels*/
data stack;
set stack;

drop smoking hasjob cgarsr6-cgarsr42 cgarsr44-cgarsr96 typen nbrpeopl
asthma_emphysema havecare hearing
inpatnights inpatstays strat psu wgt faccost instcost i_days institution
typen bc_taken bp_taken nbrrooms dwel;
run;

proc means data=stack n nolabels;
run;

proc sort data=stack;
by _mult_;
run;

proc freq data=stack;
table i*_mult_;
run;

data mult1 mult2 mult3 mult4 mult5;
set stack;
    if _mult_ = 1 then output mult1;
    if _mult_ = 2 then output mult2;
    if _mult_ = 3 then output mult3;
    if _mult_ = 4 then output mult4;
    if _mult_ = 5 then output mult5;
run;

/*impute povcat with variables common to both sets*/
options set=srclib "Insert path for \srclib" sasautos=('!SRCLIB' sasautos)
mautosource;

%impute(name=mcbs09_inst_2a, dir=Insert file path, setup=old);
%impute(name=mcbs09_inst_2b, dir=Insert file path, setup=old);
%impute(name=mcbs09_inst_2c, dir=Insert file path, setup=old);
%impute(name=mcbs09_inst_2d, dir=Insert file path, setup=old);
%impute(name=mcbs09_inst_2e, dir=Insert file path, setup=old);

data end1;
set end1;
_mult_ = 1;
run;

```

```
data end2;  
set end2;  
_mult_ = 2;  
run;
```

```
data end3;  
set end3;  
_mult_ = 3;  
run;
```

```
data end4;  
set end4;  
_mult_ = 4;  
run;
```

```
data end5;  
set end5;  
_mult_ = 5;  
run;
```

```
data combine;  
set end1 end2 end3 end4 end5;  
where i = 1;  
run;
```

```
proc sort data=combine;  
by _mult_;  
run;
```

```
proc sort data=combine;  
by baseid;  
run;
```

```
Data observed;  
Set sharedi.inst_mcbs_09;  
Run;
```

```
/*read in 09 institutionalized data and subset  
- merge back variables originally removed at beginning*/  
Data mcbs_inst_09;  
Set sharedi.inst_mcbs_09;  
keep baseid cgar1-cgar105 cgar122-cgar125 strat psu wgt  
inpatnights inpatstays faccost instcost i_days  
institution type male;  
Run;
```

```
proc sort data=combine;
by baseid;
run;

data final;
merge combine mcbs_inst_09;
by baseid;
run;

proc means data=final nlabels;
run;

proc contents data=final;
run;

proc compare base=observed compare=final briefsummary listbasevar listcompvar
novalues;
run;

options orientation=landscape;
ods rtf file='Insert file path\imp_freq_2.rtf' columns=2 ;

proc freq data=sharedi.inst_mcbs_09;
table flushot ;
run;

proc freq data=final;
table flushot*_mult_/ norow nopercnt;
run;

proc freq data=wts.mcbs_ni_i09;
table flushot;
run;

proc freq data=final;
table flushot*_mult_/ norow nopercnt;
run;

proc freq data=sharedi.inst_mcbs_09;
table povcat ;
run;
```

```

proc freq data=final;
table povcat*_mult_/ norow nopercnt;
run;

proc freq data=wts.mcbs_ni_i09;
table povcat;
run;

proc freq data=final;
table povcat*_mult_/ norow nopercnt;
run;
ods rtf close;

proc contents data=final;
run;

proc means data=final nolabels;
run;

proc compare base=observed compare=final briefsummary listbasevar listcompvar
novalues;
run;

/*****
      create final data set in appropriate folder*/

/*combine multiples of imputation */
data sharedi.Inst_MCBS_i09;
    set final;
run;

*Variable 'Type' did not merge back on for all calibrated sets, drop and add
back here;
data test;
set sharedi.Inst_MCBS_i&Y.;
drop type;
run;

data test2 (keep=baseid type);
set sharedi.inst_mcbs_&Y.;
run;

data combine2;

```

```
merge test test2;  
by baseid;  
drop nbrpeopl;  
run;
```

```
* 2009 create final data set in appropriate folder ;  
data sharedi.Inst_MCBS_i&Y.;  
  set combine2;  
run;
```

```
data sharedi.Inst_MCBS_i&Y._pure;  
set combine2;  
if pure=1;  
run;
```



```

/*****
Program: 'impute_nhanes_65_0910.sas'
Purpose: Impute within survey missing values
Data in: Insert path for input datasets
Data out: Insert path for output dataset
*****/

libname g 'Insert file path';
%include 'Insert file path\xdelete.sas';
%xdelete(_ALL_);

data nhanes;
  Set g.nhanes0910_1full;
  lweight=log(weightkg);
If age>=65;
If cgarsr16=. then do;
  diab_pill=.;
  diab_inject=.;
  diab_eye=.;
End;
drop weightkg priv_insur;
Run;

%impute(name=nhanes65_0910, dir=Insert file path, setup=old);
%putdata(name=nhanes65_0910, dir= Insert file path,mult=2, dataout=end_2);
run;
%putdata(name=nhanes65_0910, dir=Insert file path,mult=3, dataout=end_3);
run;
%putdata(name=nhanes65_0910, dir=Insert file path,mult=4, dataout=end_4);
run;
%putdata(name=nhanes65_0910, dir=Insert file path,mult=5, dataout=end_5);
run;

Data g.nhanes65_i0910;
set end_1 end_2 end_3 end_4 end_5;
  If male=0 then do;
    cgarsr11=0;
  End;
  If male=1 then do;
    hyst=0;
    cgarsr9=0;
    cgarsr10=0;
  End;
  If arthrit=1 and t_arthr=1 then cgarsr88=1;
  Else cgarsr88=0;
  If arthrit=1 and t_arthr=0 then cgarsr89=1;

```

```
Else cgarsr89=0;
If cgarsr16=0 then diab_pill=0;
If cgarsr16=0 then diab_eye=0;
If cgarsr16=0 then diab_inject=0;
If cgarsr17=2 then cgarsr17=0;
If cgarsr19=2 then cgarsr19=0;
If cgarsr50=2 then cgarsr50=0;
If cgarsr18=2 then cgarsr18=0;
weightkg=exp(lweight);
If everSmoke=0 then smokenow=0;
asthma_emphysema=max(cgarsr67, cgarsr68);
If cgarsr68=0 then asthmaattack=0;
If insur=1 then priv_insur=1;
Else priv_insur=0;
cgarsr52=max(chd, angina);

drop t_arthr lweight chd angina;
Run;
```

mcbs09.set

```
datain mcbs_ni;
dataout end1;
default categorical;
continuous lweight height age;
count inpatstays inpatnights cost nbrrooms nbrpeopl i_days;
transfer baseid wgt psu strat weightkg cgar17 cgar19 cgar50
       cgar99 cgar101;
bounds cost(>=0) nbrpeopl (>=1, <=7) nbrrooms(>=1, <=10) ;
       restrict smokenow(eversmoke=1) cgarsr11(male=1) cgarsr82(male=1)
       cgarsr10(male=0) cgarsr9(male=0)
       mammogram(MALE=0) psa1yr(MALE=1) pap_smear(MALE=0) ;
iterations 5;
multiples 5;
minrsqd 0.1;
seed 2015;
run;
```

mcbs09_inst_2a.set

```
datain mult1; *input data set;  
dataout end1; *output after imputation of first multiple;
```

```
    categorical cgarsr43 died male race didserv marital s ed5 flushot pneushot  
healthstat hearingaid ever smoke  
smokenow dif_lif t dif_stoop dif_walk prb_dres prb_eat priv_insur psa1yr pap_smear  
hyst mammogram povcat pure  
comphealth;
```

```
transfer baseid hear_inst _mult_ type mcadv_months mcab_months_3grp i;
```

```
mi nrsqd 0.0001;  
iterations 5;  
multiples 1;  
seed 2015;  
run;
```

mcbs09_inst_2b.set

```
datain mult2; *input data set;  
dataout end2; *output after imputation of first multiple;
```

```
    categorical cgarsr43 died male race didserv marital s ed5 flushot pneushot  
healthstat hearingaid ever smoke  
smokenow dif_lif t dif_stoop dif_walk prb_dres prb_eat priv_insur psa1yr pap_smear  
hyst mammogram povcat pure  
comphealth;
```

```
transfer baseid hear_inst _mult_ type mcadv_months mcab_months_3grp i;
```

```
mi nrsqd 0.0001;  
iterations 5;  
multiples 1;  
seed 2015;  
run;
```

mcbs09_inst_2c.set

```
datain mult3; *input data set;  
dataout end3; *output after imputation of first multiple;
```

```
    categorical cgarsr43 died male race didserv marital s ed5 flushot pneushot  
healthstat hearingaid ever smoke  
smokenow dif_lifft dif_stoop dif_walk prb_dres prb_eat priv_insur psa1yr pap_smear  
hyst mammogram povcat pure  
comphealth;
```

```
transfer baseid hear_inst _mult_ type mcadv_months mcab_months_3grp i;
```

```
mi nrsqd 0.0001;  
iterations 5;  
multiples 1;  
seed 2015;  
run;
```

mcbs09_inst_2d.set

```
datain mult4; *input data set;  
dataout end4; *output after imputation of first multiple;
```

```
    categorical cgarsr43 died male race didserv marital s ed5 flushot pneushot  
healthstat hearingaid ever smoke  
smokenow dif_lif t dif_stoop dif_walk prb_dres prb_eat priv_insur psa1yr pap_smear  
hyst mammogram povcat pure  
comphealth;
```

```
transfer baseid hear_inst _mult_ type mcadv_months mcab_months_3grp i;
```

```
mi nrsqd 0.0001;  
iterations 5;  
multiples 1;  
seed 2015;  
run;
```

mcbs09_inst_2e.set

```
datain mult5; *input data set;  
dataout end5; *output after imputation of first multiple;
```

```
    categorical cgarsr43 died male race didserv marital s ed5 flushot pneushot  
healthstat hearingaid ever smoke  
smokenow dif_lif t dif_stoop dif_walk prb_dres prb_eat priv_insur psa1yr pap_smear  
hyst mammogram povcat pure  
comphealth;
```

```
transfer baseid hear_inst _mult_ type mcadv_months mcab_months_3grp i;
```

```
mi nrsqd 0.0001;  
iterations 5;  
multiples 1;  
seed 2015;  
run;
```


nhanes65_0910.set

```
title Multiple imputation;
datain nhanes;
dataout end_1;
default categorical;
continuous lweight height;
count inpatstays nbrrooms nbrpeopl;
transfer baseid wgt psu strat;
bounds inpatstays(<=365) nbrpeopl (>0, <=7) nbrrooms(>0, <=10);
restrict smokenow(eversmoke=1) cgarsr11(male=1) cgarsr10(male=0)
cgarsr9(male=0) t_arthr(arthrit=1)
diab_inject(cgarsr16=1) diab_pill(cgarsr16=1) diab_eye(cgarsr16=1)
asthmaattack(cgarsr68=1)
cgarsr19(cgarsr18=0) cgarsr50(cgarsr49=0) cgarsr17(cgarsr16=0)
cgarsr18(bc_taken<5) ;
mi nrsqd 0.025;
iterations 5;
multiples 5;
seed 2015;
run;
```