

HINTS 4 Cycle 2 History Document

Iteration One (finalized on May 26, 2013)

Data Editing

The following variables were identified to contain invalid or unusual values. Those values were replaced with negative value of -4, “Unreadable or Non-conforming numeric response”, negative value of -9, “Missing data (Not Ascertained)” or reasonable regular values.

MailHHAdults: 2. Including yourself, how many people age 18 or older live in this household?

Three respondents had the value of 0, which was replaced by 1.

Age: 01. (L1.) What is your age?

One respondent had age of 176, which was replaced by -4.

Height_Feet: F5. (E4.) About how tall are you without shoes? Feet:

One respondent had the feet of 0; two respondents had the feet of 2; one respondent had the feet of 3; two respondents had the feet of 8; and two respondents had the feet of 9. All those values were replaced by -4.

Height_Inches: F5. (E4.) About how tall are you without shoes? Inches:

Two hundred and thirty two respondents had inches of -9, “Missing data (Not Ascertained)”, which was replaced by 0. The variable Height_Feet is recommended for determining validity of full height.

Weight: F6. (E5.) About how much do you weigh, in pounds, without shoes?

One respondent had weight of 0; one respondent had weight of 15; one respondent had weight of 16; one respondent had weight of 23; and one respondent had weight of 35. All their weights were replaced by -4.

YearCameToUSA: O8. (L8.) In what year did you come to live in the United States?

Thirteen respondents had values between 1 and 82, which were replaced by -4.

SelfAge: O13. (L13.) Starting with yourself, please mark the gender, and write in the age and month of birth for each adult 18 years of age or older living at this address. Self: Age.

One respondent had age of 1; one respondent had age of 16; and one respondent had age of 17. All their ages were replaced by -4.

SelfMOB: O13. (L13.) Starting with yourself, please mark the gender, and write in the age and month of birth for each adult 18 years of age or older living at this address. Self: Month Born.

Three respondents had self month of birth of 0; and seventy six respondents had self month of birth greater than 12. All their self month of birth were replaced by -4.

HHAdultAge2: O13. (L13.) Starting with yourself, please mark the gender, and write in the age and month of birth for each adult 18 years of age or older living at this address. Adult 2: Age

Forty three respondents had Adult 2 age between 1 and 17. All their ages were replaced by -4.

HHAdultMOB2: O13. (L13.) Starting with yourself, please mark the gender, and write in the age and month of birth for each adult 18 years of age or older living at this address. Adult 2: Month Born

Three respondents had Adult 2 month of birth of 0; and fifty respondents had Adult 2 month of birth greater than 12. All their months of birth were replaced by -4.

HHAdultAge3: O13. (L13.) Starting with yourself, please mark the gender, and write in the age and month of birth for each adult 18 years of age or older living at this address. Adult 3: Age

One hundred and seventy eight respondents had Adult 3 age between 0 and 17. All their ages were replaced by -4.

HHAdultMOB3: 013. (L13.) Starting with yourself, please mark the gender, and write in the age and month of birth for each adult 18 years of age or older living at this address. Adult 3: Month Born

Three respondents had Adult 3 month of birth of 0; and twenty one respondents had Adult 3 month of birth greater than 12. All their months of birth were replaced by -4.

HHAdultAge4: 013. (L13.) Starting with yourself, please mark the gender, and write in the age and month of birth for each adult 18 years of age or older living at this address. Adult 4: Age

One hundred and twenty four respondents had Adult 4 age between 0 and 17. All their ages were replaced by -4.

HHAdultMOB4: 013. (L13.) Starting with yourself, please mark the gender, and write in the age and month of birth for each adult 18 years of age or older living at this address. Adult 4: Month Born

Six respondents had Adult 4 month of birth greater than 12. All their months of birth were replaced by -4.

HHAdultAge5: 013. (L13.) Starting with yourself, please mark the gender, and write in the age and month of birth for each adult 18 years of age or older living at this address. Adult 5: Age

Fifty five respondents had Adult 5 age between 0 and 17. All their ages were replaced by -4.

R_HHAdults: Reconciled number of adults in household

Four respondents had missing values (.), which were replaced by -9.

Standard Recode

Standard recode/derived variables are listed below.

AgeGrpA: -->AgeGrpA. (Age Recode -- 4 Levels)

The Age variable was re-coded into 4 categories: 18-34; 35-39; 40-44; 45+. The original negative values were carried over.

AgeGrpB: -->AgeGrpB. (Age Recode -- 5 Levels)

The Age variable was re-coded into 5 categories: 18-34; 35-49; 50-64; 65-74; 75+. The original negative values were carried over.

EducA: -->EducA. What is the highest level of school you completed? (Education Recode -- 4 Levels)

The Education variable was re-coded into 4 categories: Less than High School; High School Graduate; Some College; College Graduate or More. The original negative values were carried over.

EducB:-->EducB. What is the highest level of school you completed? (Education Recode -- 5 Levels)

The Education variable was re-coded into 5 categories: Less than High School; High School Graduate; Some College; Bachelor's Degree; Post-Baccalaureate Degree. The original negative values were carried over.

RaceEthn:-->Race/Ethnicity. (Hisp_Cat and Race_Cat2 Recode -- 7 Levels)

The RaceEthn was created with Hisp_Cat and Race_Cat2 variables. The RaceEthn has 7 categories: Hispanic; Non-Hispanic White; Non-Hispanic Black or African American; Non-Hispanic American Indian or Alaska Native; Non-Hispanic Asian; Non-Hispanic Native Hawaiian or other Pacific Islander; Non-Hispanic Multiple Races Mentioned. If Hisp_Cat had value of 10, "Not Hispanic", and Race_Cat2 had value of -9, "Missing data (Not Ascertained)", the RaceEthn was assigned with value of -9. The RaceEthn was assigned with value of -9 if Hispanic=-9.

HHInc:-->HHInc. What is your {combined} annual household income? (IncomeRanges Recode -- 5 Levels)

The IncomeRanges variable was re-coded into 5 categories: Less than \$20,000; \$20,000 to < \$35,000; \$35,000 to < \$50,000; \$50,000 to < \$75,000; \$75,000 or more. The original negative values were carried over.

BMI:-->BMI. Body Mass Index (Weight*703)/(Height in inches2)**

The BMI variable was created with weight in pounds and height in inches. If feet of height or weight had value of -9, "Missing data (Not Ascertained)" but neither had value of -4, "Unreadable or Non-conforming numeric response", the BMI was assigned to -9. If feet of height or weight had value of -4, "Unreadable or Non-conforming numeric response", the BMI was assigned to -4.

AgeDX:-->AgeDX. At what age were you diagnosed with cancer? (WhenDiagnosedCancer Recode)

The variable AgeDX is a copy of variable WhenDiagnosedCancer.

TimeSinceDX:-->TimeSinceDX. How long ago were you diagnosed with cancer? (WhenDiagnosedCancer Age Recode)

The variable TimeSinceDX was created with EverHadCancer, WhenDiagnosedCancer and Age variables. The variable TimeSinceDX has 4 categories: Less than 1 Year since DX; 2-5 Years since DX; 6-10 Years since DX; 11+ Years since DX. If the variable EverHadCancer had value of 1 and either Age or WhenDiagnosedCancer had value of -9, "Missing data (Not Ascertained)", the TimeSinceDX was assigned to -9. If the variable EverHadCancer had value of 1 and WhenDiagnosedCancer is greater than Age, the TimeSinceDX was assigned to -4. If the variable EverHadCancer had value of -9, the TimeSinceDX was assigned to -6. If the variable EverHadCancer had value of 2 and WhenDiagnosedCancer had value of -2, the TimeSinceDX was assigned to -4. If the variable EverHadCancer had value of 2 and WhenDiagnosedCancer had value of -1, the TimeSinceDX was assigned to -1. .

smokeStat:-->smokeStat. Smoking Status (Smoke100 and SmokeNow Recode)

The variable smokeStat was created with Smoke100 and SmokeNow variables. The variable smokeStat has 3 categories: Current; Former; Never. If Smoke100 had value of 1 and SmokeNow had value of -5, "Multiple responses selected in error", the smokeStat was assigned to -4. If Smoke100 had value of 1 and SmokeNow had value of -9, "Missing data (Not Ascertained)", the smokeStat was assigned to -9. If Smoke100 had value of -9, "Missing data (Not Ascertained)", the smokeStat was assigned to -6.

PHQ4:-->PHQ4. PHQ-4 total score (LittleInterest Hopeless Nervous Worrying Combined)

The variable PHQ4 was created by summing the reversed LittleInterest, Hopeless, Nervous and Worrying variables. If any variable had value of -5, "Multiple responses selected in error", the PHQ4 was assigned to -4. If any variable had value of -9, "Missing data (Not Ascertained)", and no -5, the PHQ4 was assigned to -9.

Label Editing

The labels of following twenty five variables were edited: ActiveDutyArmedForces, AverageDailyTVGames, CancerCheckups, CancerConcernedQuality, CancerFrustrated, CancerLotOfEffort, CancerTooHardUnderstand, ChanceAskQuestions, ElectCigLessHarm, EverHadPSATest, ExplainedClearly, FeelingsAddressed, Genetics_Diabetes, Genetics_HighBP, HeardGeneticTest, HelpUncertainty, InvolvedDecisions, RegExercise_Enjoyment, RegExercise_Guilt, TimesModerateExercise, TimesStrengthTraining, UnderstoodNextSteps, EverTalkedHPVShot, HeardOfClinicalTrial and SpentEnoughTime.

Variable Exclusion

The following variables were excluded from the public data file. They are: FIPST, FIPS, HealthInsurance_I, HealthInsurance_IFlag, GenderC_I, Gender_IFlag, EverHadCancer_I, Cancer_IFlag, Age_I, Age_IFlag, MaritalStatus_I, Marital_IFlag, Education_I, Education_IFlag, Hisp_Cat_I, Hisp_IFlag, Race_Cat2_I, Race_IFlag, WhereSeekHealthInfo_IMP, WhereSeekHealthInfo_IFlag, StrongNeedCancerInfo_IMP and StrongNeedCancerInfo_IFlag.

Format Editing

Format Added for Existing Variable YearCameToUSA

The format YearCameToUSA was created and assigned to variable YearCameToUSA because the values of YearCameToUSA were edited.

Formats Added for Standard Recode Variables

The formats AgeGrpA, AgeGrpB, Educa, EducB, RaceEthn, HHInc, BMI, AgeDX, TimeSinceDX, smokeStat and PHQ4Total were created and assigned to the variables AgeGrpA, AgeGrpB, Educa, EducB, RaceEthn, HHInc, BMI, AgeDX, TimeSinceDX, smokeStat and PHQ4, respectively.

Formats Modified for Certain Variables

All ranges in formats were excluded, which allows the raw values to be accessed. The modified formats are: Age, Average, Childre, HeightF, Height_, HHAdulF, HHAdulG, HHAdulH, HHAdulI, HHAdulJ, HHAdulK, HHAdulL, HHAdulT, HowLonF, HowLonG, MailHHA, MailSuf, MailSuR, MultiOc, SelfAge, SelfMOB, SleepWe, SleepWf, SleepWg, SleepWo, TotalHo, Weight, WhenDia, YearCam, \$CaOthef, \$InternH, \$InternI, \$Occupat, \$StrongN and \$WhereSG.

Formats Disassociated or Re-associated for Certain Variables

The variable R_HHAdults was formatted with format Childre. The variable MultiOcc was unformatted.

Imputation of Income Variable

The income variable (IncomeRanges) has relatively higher percentage (13.2% for un-weighted percentage or 10.5% for weighted percentage) of missing values. This variable was imputed via PROC HOTDECK in SUDAAN. The imputation class variables are: Education (O6), RaceEthn (standard recode), RentOrOwn (O15), BornInUSA (O7) and SpeakEnglish (O9). Since the variable SpeakEnglish was asked for people who were born outside USA (BornInUSA=2), the variable SpeakEnglish was declared after

BornInUSA in imputation class statement. The copy variables of the imputation class variables and income variable were created, where the missing values were appropriately coded. The copy variables are used for the imputation. The imputed values were saved in a new variable IncomeRanges_IMP.

SAS Code for Data Editing

```
* Recode MailHHAdults of 0 to 1
*;
    if MailHHAdults in (0) then
        MailHHAdults = 1;

* Recode Height_Feet of 0, 2, 3, 8 and 9 to Unreadable or Non-
conforming numeric response
    if Height_Feet in (0, 2, 3, 8, 9) then
        do;
            Height_Feet = -4;
        end;

* Recode Height_Inches of -9 to 0
*;
    if Height_Inches in (-9) then
        Height_Inches = 0;

* Recode Weight of 0, 15, 16, 23 and 35 to Unreadable or Non-
conforming numeric response
    if Weight in (0, 15, 16, 23, 35) then
        Weight = -4;

* Recode Age of 176 to Unreadable or Non-conforming numeric response
*;
    if Age in (176) then
        Age = -4;

* Recode YearCameToUSA between 1 and 82 to Unreadable or Non-
conforming numeric response
    if 1 <= YearCameToUSA <= 82 then
        YearCameToUSA = -4;
```

```

* Recode SelfAge between 0 and 17 to Unreadable or Non-conforming
numeric response          *;
    if 0 <= SelfAge < 18 then
        SelfAge = -4;

* Recode SelfMOB of 0 and SelfMOB > 12 to Unreadable or Non-conforming
numeric response          *;
    if SelfMOB in (0) then
        SelfMOB = -4;
    else if SelfMOB > 12 then
        SelfMOB = -4;

* Recode HHAdultAge2-HHAdultAge5 between 0 and 17 to Unreadable or
Non-conforming numeric response*;
* We leave HHAdultAge2 of 110 unchanged
*;
    array a(*) HHAdultAge2-HHAdultAge5;

    do I = 1 to dim(a);
        if 0 <= a(I) < 18 then
            a(I) = -4;
    end;

* Recode HHAdultMOB2-HHAdultMOB5 of 0 to Unreadable or Non-conforming
numeric response          *;
* Recode HHAdultMOB2-HHAdultMOB5 > 12 to Unreadable or Non-conforming
numeric response          *;
    array b(*) HHAdultMOB2-HHAdultMOB5;

    do I = 1 to dim(b);
        if b(I) = 0 then
            b(I) = -4;
        else if b(I) > 12 then
            b(I) = -4;
    end;

    drop I;

* Recode R_HHAdults of missing (.) to -9
*;
    if missing(R_HHAdults) = 1 then
        R_HHAdults = -9;

```


SAS Code for Standard Recode

```
if 18 <= Age <= 34 then
    AgeGrpA = 1;
else if 35 <= Age <= 39 then
    AgeGrpA = 2;
else if 40 <= Age <= 44 then
    AgeGrpA = 3;
else if 45 <= Age then
    AgeGrpA = 4;
else if Age in (-4, -9) then
    AgeGrpA = Age;
label AgeGrpA = '-->AgeGrpA. (Age Recode -- 4 Levels)';

if 18 <= Age <= 34 then
    AgeGrpB = 1;
else if 35 <= Age <= 49 then
    AgeGrpB = 2;
else if 50 <= Age <= 64 then
    AgeGrpB = 3;
else if 65 <= Age <= 74 then
    AgeGrpB = 4;
else if 75 <= Age then
    AgeGrpB = 5;
else if Age in (-4, -9) then
    AgeGrpB = Age;
label AgeGrpB = '-->AgeGrpB. (Age Recode -- 5 Levels)';

if Education in (1, 2) then
    EducA = 1;
else if Education in (3) then
    EducA = 2;
else if Education in (4, 5) then
    EducA = 3;
else if Education in (6, 7) then
    EducA = 4;
else if Education in (-9) then
    EducA = Education;
label EducA = '-->EducA. What is the highest level of school you
completed? (Education Recode -- 4 Levels)';

if Education in (1, 2) then
    EducB = 1;
else if Education in (3) then
    EducB = 2;
else if Education in (4, 5) then
    EducB = 3;
else if Education in (6) then
    EducB = 4;
```

```

else if Education in (7) then
    EducB = 5;
else if Education in (-9) then
    EducB = Education;
label EducB = '-->EducB. What is the highest level of school you
completed? (Education Recode -- 5 Levels)';

if Hisp_Cat in (21, 22, 23, 24, 25) then
    RaceEthn = 1;
else if Hisp_Cat in (10) then
    do;
        if Race_Cat2 in (11) then
            RaceEthn = 2;
        else if Race_Cat2 in (12) then
            RaceEthn = 3;
        else if Race_Cat2 in (14) then
            RaceEthn = 4;
        else if Race_Cat2 in (31, 32, 33, 34, 35, 36, 37) then
            RaceEthn = 5;
        else if Race_Cat2 in (51, 52, 53, 54) then
            RaceEthn = 6;
        else if Race_Cat2 in (16) then
            RaceEthn = 7;
        else if Race_Cat2 in (-9) then
            RaceEthn = -9;
    end;
else if Hisp_Cat in (-9) then
    do;
        RaceEthn = -9;
    end;
label RaceEthn = '-->Race/Ethnicity. (Hisp_Cat and Race_Cat2
Recode -- 7 Levels)';

if IncomeRanges in (1, 2, 3) then
    HHInc = 1;
else if IncomeRanges in (4) then
    HHInc = 2;
else if IncomeRanges in (5) then
    HHInc = 3;
else if IncomeRanges in (6) then
    HHInc = 4;
else if IncomeRanges in (7, 8, 9) then
    HHInc = 5;
else if IncomeRanges in (-9) then
    HHInc = IncomeRanges;
label HHInc = '-->HHInc. What is your {combined} annual
household income? (IncomeRanges Recode -- 5 Levels)';

if Height_Feet >= 0 and Height_Inches >= 0 and Weight > 0 then
    BMI = (Weight * 703) / ((Height_Feet * 12 +
Height_Inches)**2);

```

```

    else if (Height_Feet = -9 and Weight > 0 ) or (Height_Feet >=0
and Weight = -9) or (Height_Feet = -9 and Weight = -9) then
        BMI = -9;
    else if Height_Feet = -4 or Weight = -4 then
        BMI = -4;
    label BMI = '-->BMI.  Body Mass Index (Weight*703)/(Height in
inches**2)';
    if BMI not in (-4, -9) then
        BMI = round(BMI, 0.1);

    AgeDX = WhenDiagnosedCancer;
    label AgeDX = '-->AgeDX. At what age were you diagnosed with
cancer? (WhenDiagnosedCancer Recode)';

    if EverHadCancer in (1) then
        do;
            if Age in (-9) then
                TimeSinceDX = Age;
            else if WhenDiagnosedCancer in (-9) and Age >= 18 then
                TimeSinceDX = WhenDiagnosedCancer;
            else if WhenDiagnosedCancer >= 0 and Age >= 18 then
                do;
                    if 0 <= (Age - WhenDiagnosedCancer) <= 1 then
                        TimeSinceDX = 1;
                    else if 2 <= (Age - WhenDiagnosedCancer) <= 5 then
                        TimeSinceDX = 2;
                    else if 6 <= (Age - WhenDiagnosedCancer) <= 10 then
                        TimeSinceDX = 3;
                    else if 11 <= (Age - WhenDiagnosedCancer) then
                        TimeSinceDX = 4;
                    else if (Age - WhenDiagnosedCancer) < 0 then
                        TimeSinceDX = -4;
                end;
            end;
        else if EverHadCancer in (-9) then
            TimeSinceDX = -6;
        else if EverHadCancer in (2) then
            do;
                if WhenDiagnosedCancer in (-1) then
                    TimeSinceDX = WhenDiagnosedCancer;
                else if WhenDiagnosedCancer in (-2) then
                    TimeSinceDX = -4;
            end;
        label TimeSinceDX = '-->TimeSinceDX.  How long ago were you
diagnosed with cancer? (WhenDiagnosedCancer Age Recode)';

    if Smoke100 in (1) then
        do;
            if SmokeNow in (1, 2) then
                smokeStat = 1;
            else if SmokeNow in (3) then
                smokeStat = 2;

```

```

        else if SmokeNow in (-5) then
            smokeStat = -4;
        else if SmokeNow in (-9) then
            smokeStat = -9;
    end;
else if Smoke100 in (2) then
    do;
        smokeStat = 3;
    end;
else if Smoke100 in (-9) then
    smokeStat = -6;
label smokeStat = '-->smokeStat. Smoking Status (Smoke100 and
SmokeNow Recode)';

array a(*) LittleInterest Hopeless Nervous Worrying;

PHQ4 = 0;
if a(1) = -5 or a(2) = -5 or a(3) = -5 or a(4) = -5 then
    PHQ4 = -4;
else
    do;
        if a(1) = -9 or a(2) = -9 or a(3) = -9 or a(4) = -9
then
            PHQ4 = -9;
        else
            do;
                do I = 1 to dim(a);
                    if PHQ4 not in (-5, -9) and a(I) in
(1, 2, 3, 4) then
                        PHQ4 = PHQ4 + (4 - a(I));
                end;
            end;
        end;
label PHQ4 = '-->PHQ4. PHQ-4 total score (LittleInterest Hopeless
Nervous Worrying Combined)';

drop I;

```

SAS Code for Label Editing

```

label

    ActiveDutyArmedForces = "03. (L3.) Have you ever served on
active duty in the U.S. Armed Forces, military Reserves or National
Guard?"

```

AverageDailyTVGames = "H6. (F8.) Over the past 30 days, how many hours per day, on average, did you sit and watch TV or movies, surf the web, or play computer games?"

CancerCheckups = "M8. Have you ever received instructions from a doctor or other health care professional about where you should return...after completing your cancer treatment?"

CancerConcernedQuality = "A6c. Based on the results of your most recent search for information about cancer how much do you agree or disagree with each of the following statements? Concerned about Quality."

CancerFrustrated = "A6b. Based on the results of your most recent search for information about cancer, how much do you agree or disagree with each of the following statements? Frustrated."

CancerLotOfEffort = "A6a. Based on the results of your most recent search for information about cancer, how much do you agree or disagree with each of the following statements? Took effort."

CancerTooHardUnderstand = "A6d. Based on the results of your most recent search for information about cancer, how much do you agree or disagree with each of the following statements? Hard to understand."

ChanceAskQuestions = "C5a. In the past 12 months, how often did your health professional: Give you the chance to ask all the health-related questions you had? "

ElectCigLessHarm = "J10. Compared to smoking cigarettes, would you say that electronic cigarettes are..."

EverHadPSATest = "L8. (I4.) Have you ever had PSA test?"

ExplainedClearly = "C5e. In the past 12 months, how often did your health professional: Explain things in a way you could understand?"

FeelingsAddressed = "C5b. In the past 12 months, how often did your health professional: Give the attention you needed to your feelings and emotions?"

Genetics_Diabetes = "E4a. How much do you think genetics determine whether or not a person will develop each of the following conditions? Diabetes or high blood sugar?"

Genetics_HighBP = "E4d. How much do you think genetics determine whether or not a person will develop each of the following conditions? High blood pressure or hypertension?"

HeardGeneticTest = "E2. (A11.) Genetic tests that analyze your DNA, diet and lifestyle are currently being marketed by companies directly to consumers. Have you ever heard or read about these tests?"

HelpUncertainty = "C5g. In the past 12 months, how often did your health professional: Help you deal with feelings of uncertainty about your health or health care?"

InvolvedDecisions = "C5c. In the past 12 months, how often did your health professional: Involve you in decisions about your health care as much as you wanted?"

RegExercise_Enjoyment = "H5d. People start or continue exercising regularly for lots of reasons. How much does getting enjoyment reflect why you would start or continue?"

RegExercise_Guilt = "H5c. People start or continue exercising regularly for lots of reasons. How much does feeling guilty when you skip reflect why you would start or continue?"

TimesModerateExercise = "H1. (F5.) In a typical week, how many days do you do any physical activity of at least moderate intensity?"

TimesStrengthTraining = "H3. (F7.) In a typical week, how many days do you do leisure-time physical activities specifically designed to strengthen your muscles?"

UnderstoodNextSteps = "C5d. In the past 12 months how often did your health professional: Make sure you understood the things you needed to do to take care of your health?"

EverTalkedHPVShot = "L1. Has a doctor or other health care professional ever talked with you about the HPV shot or vaccine?"

HeardOfClinicalTrial = "E1. Clinical trials are research studies that involve people. Have you ever heard of a clinical trial?"

SpentEnoughTime = "C5f. In the past 12 months, how often did your health professional: Spend enough time with you?"

;

SAS Code for Variable Exclusion

```
drop FIPST FIPS
    HealthInsurance_I HealthInsurance_IFlag
    GenderC_I Gender_IFlag
    EverHadCancer_I Cancer_IFlag
    Age_I Age_IFlag
    MaritalStatus_I Marital_IFlag
    Education_I Education_IFlag
    Hisp_Cat_I Hisp_IFlag
    Race_Cat2_I Race_IFlag
    WhereSeekHealthInfo_IMP WhereSeekHealthInfo_IFlag
    StrongNeedCancerInfo_IMP StrongNeedCancerInfo_IFlag
;
```

SAS Code for Format Editing

SAS Code for Format Added for Existing Variable YearCameToUSA

```
value YearCameToUSA -1 = 'Inapplicable, coded 1 in BornInUSA'
                    -2 = 'Question answered in error
(Commission Error)'
                    -4 = 'Unreadable or Nonconforming
Numeric Response'
                    -6 = 'Missing data (Filter Missing)'
                    -9 = 'Missing data (Not Ascertained)'
                    ;
```

SAS Code for Formats Added for Standard Recode Variables

```
value AgeGrpA 1 = '18-34'
              2 = '35-39'
              3 = '40-44'
              4 = '45+'
              -4 = 'Unreadable or Nonconforming Numeric
Response'
              -9 = 'Missing Data (Not Ascertained)'
              ;
```

```
value AgeGrpB 1 = '18-34'
              2 = '35-49'
              3 = '50-64'
              4 = '65-74'
              5 = '75+'
              -4 = 'Unreadable or Nonconforming Numeric
Response'
              -9 = 'Missing Data (Not Ascertained)'
              ;
```

```
value EducA 1 = 'Less than High School'
              2 = 'High School Graduate'
              3 = 'Some College'
              4 = 'College Graduate or More'
              -9 = 'Missing Data (Not Ascertained)'
              ;
```

```
value EducB 1 = 'Less than High School'
              2 = 'High School Graduate'
```



```

3 = 'Some College'
4 = 'Bachelor's Degree'
5 = 'Post-Baccalaureate Degree'
-9 = 'Missing Data (Not Ascertained)'
;

value RaceEthn 1 = 'Hispanic'
2 = 'Non-Hispanic White'
3 = 'Non-Hispanic Black or African
American'
4 = 'Non-Hispanic American Indian or Alaska
Native'
5 = 'Non-Hispanic Asian'
6 = 'Non-Hispanic Native Hawaiian or other
Pacific Islander'
7 = 'Non-Hispanic Multiple Races Mentioned'
-4 = 'Unreadable or Nonconforming Numeric
Response'
-9 = 'Missing Data (Not Ascertained)'
;

value HHInc 1 = 'Less than $20,000'
2 = '$20,000 to < $35,000'
3 = '$35,000 to < $50,000'
4 = '$50,000 to < $75,000'
5 = '$75,000 or More'
-5 = 'Multiple Responses Selected in Error'
-9 = 'Missing Data (Not Ascertained)'
;

value BMI -4 = 'Unreadable or Nonconforming Numeric Response'
-9 = 'Missing Data (Not Ascertained)'
;

value AgeDX -1 = 'Inapplicable, coded 2 in EverHadCancer'
-2 = 'Question Answered in Error (Commission
Error)'
-6 = 'Missing Data (Filter Missing)'
-9 = 'Missing Data (Not Ascertained)'
;

value TimeSinceDX 1 = 'Less than 1 Yr Since DX'
2 = '2-5 Yrs Since DX'
3 = '6-10 Yrs Since DX'
4 = '11+ Yrs Since DX'
-1 = 'Inapplicable, coded 2 in
EverHadCancer'
-4 = 'Unreadable or Nonconforming Numeric
Response'
-6 = 'Missing Data (Filter Missing), coded
-9 in EverHadCancer'
-9 = 'Missing Data (Not Ascertained)'

```

```

;

value smokeStat 1 = 'Current'
                2 = 'Former'
                3 = 'Never'
               -4 = 'Unreadable or Nonconforming Numeric
Response'
               -6 = 'Missing Data (Filter Missing), coded
-9 in Smoke100'
               -9 = 'Missing Data (Not Ascertained)'
;

value PHQ4Total -4 = 'Unreadable or Nonconforming Numeric
Response'
               -5 = 'Multiple Responses Selected in Error'
               -9 = 'Missing Data (Not Ascertained)'
;

format          AgeGrpA AgeGrpA.
                AgeGrpB AgeGrpB.
                EducA  EducA.
                EducB  EducB.
                RaceEthn RaceEthn.
                HHInc  HHInc.
                BMI    BMI.
                AgeDX  AgeDX.
                TimeSinceDX TimeSinceDX.
                smokeStat smokeStat.
                PHQ4    PHQ4Total.
                YearCameToUSA YearCameToUSA.
;

```

SAS Code for Formats Modified for Certain Variables

```

* Edit format data file
*;
data HINTS4CYCLE2FORMATS;
  set HINTS4CYCLE2FORMATS;

  if Start ^= End then
    delete;

  if Label = "Out of range" then
    delete;
run;

```

SAS Code for Formats Disassociated or Re-associated for Certain Variables

```
format R_HHAdults Childre.;
```

```
format MultiOcc;
```

SAS Code for Imputation of Income Variable

```
* Impute IncomeRanges via PROC HOTDECK
*;
data HINTS4CYCLE2;
    set HINTS4CYCLE2;

    COPY_Education = Education;
    if COPY_Education in (-9) then
        COPY_Education = .;

    COPY_RaceEthn = RaceEthn;
    if COPY_RaceEthn in (-9) then
        COPY_RaceEthn = .;

    COPY_RentOrOwn = RentOrOwn;
    if COPY_RentOrOwn in (-5, -9) then
        COPY_RentOrOwn = .;

    COPY_SpeakEnglish = SpeakEnglish;
    if COPY_SpeakEnglish in (-1, -2, -5, -6, -9) then
        COPY_SpeakEnglish = .;

    COPY_BornInUSA = BornInUSA;
    if COPY_BornInUSA in (-9) then
        COPY_BornInUSA = .;

    COPY_IncomeRanges = IncomeRanges;
    if COPY_IncomeRanges in (-9) then
        COPY_IncomeRanges = .;

    ID = _N_;

    format COPY_Education Educati. COPY_RaceEthn RaceEthn.
    COPY_RentOrOwn RentOrO.
               COPY_SpeakEnglish SpeakEn. COPY_BornInUSA BornInU.;
run;
```

```

proc freq data=HINTS4CYCLE2;
    tables COPY_Education*Education / list missing;
    tables COPY_RaceEthn*RaceEthn / list missing;
    tables COPY_RentOrOwn*RentOrOwn / list missing;
    tables COPY_SpeakEnglish*SpeakEnglish / list missing;
    tables COPY_BornInUSA*BornInUSA / list missing;
    tables COPY_IncomeRanges*IncomeRanges / list missing;
run;

proc hotdeck data=HINTS4CYCLE2 notsorted;
    weight person_finwt0;
    impvar COPY_IncomeRanges;
    impby COPY_Education COPY_RaceEthn COPY_RentOrOwn COPY_BornInUSA
COPY_SpeakEnglish;
    impname COPY_IncomeRanges="IncomeRanges_IMP";
    impid ID;
    output IMPID IMPBY IMPUTEVAL / filename=imputel replace;
run;

proc freq data=imputel;
    tables IncomeRanges_IMP / missing;
run;

proc contents data=imputel;
run;

proc sort data=HINTS4CYCLE2;
    by ID;
run;

proc sort data=imputel (keep=ID IncomeRanges_IMP);
    by ID;
run;

data HINTS4CYCLE2;
    merge HINTS4CYCLE2 (in=A) imputel (in=B);
    by ID;

    if A = 1 and B = 1;
run;

data _null_;
    set HINTS4CYCLE2;

    if IncomeRanges not in (-9) and COPY_IncomeRanges ^=
IncomeRanges_IMP then
        put ID IncomeRanges COPY_IncomeRanges IncomeRanges_IMP;
run;

data HINTS4CYCLE2;
    set HINTS4CYCLE2;

```

```
    if missing(IncomeRanges_IMP) = 1 then
        IncomeRanges_IMP = IncomeRanges;
    label IncomeRanges_IMP = '-->IncomeRanges_IMP.  Imputed
IncomeRanges variable via PROC HOTDECK in SUDAAN';
    format IncomeRanges_IMP IncomeR.;

    drop COPY_Education COPY_RaceEthn COPY_RentOrOwn
COPY_SpeakEnglish COPY_BornInUSA
        ID
        COPY_IncomeRanges;

    format MultiOcc;
run;
```