HINTS 6 History Document

April 2024

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.

PROMPT_O: Indicator of whether web respondent received prompt interventions (See Methods Report):

Dropped PROMPT_O variable.

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

One respondent had a value of 0 adults in household, which was replaced by 1.

R_HHAdults: Reconciled number of adults in household

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

Totalhousehold: P12. Including yourself, how many people live in your household?

Seventy-two respondents had Totalhousehold of 0, which were replaced by 1.

AverageTimeSitting: K4. During the past 7 days, how much time did you spend sitting on a typical day at home or at work?

Forty-six respondents had an outlying value of 21 hours or greater for AverageTimeSitting, which was replaced by -4.

HowLongModerateExerciseMinutes: M2. On the days that you do any physical activity or exercise of at least moderate intensity, how long do you typically do these activities?:

One respondent had an outlying value of 999, which was replaced by -4, "Unreadable or Non-conforming numeric response".

Variable Removal

The variables HIGHSPANLI, HISPSURNAME, HISP_HH, PersonID, and QDISP were dropped from the data.

Standard Recode

Standard recode/derived variables are listed below.

AgeGrpA: 4 Level Age Categories Version A (Derived from Age)

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

AgeGrpB: 5 Level Age Categories Version B (Derived from Age)

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: What is the highest level of school you completed? 4 Levels (Derived from Education)

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: What is the highest level of school you completed? 5 Levels (Derived from Education)

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. 7 Levels (Derived from Hisp_Cat and Race_Cat2)

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)", then RaceEthn was assigned with value of -9. If Hisp_Cat had value of 10, "Not Hispanic", and Race_Cat2 had value of -7, "Missing data (Web partial - Question Never Seen)", then RaceEthn was assigned with value of -7. RaceEthn was assigned with value of -9 if Hisp_Cat =-9. RaceEthn was assigned a value of -7 if Hisp_Cat=-7.

RaceEthn5: Race/Ethnicity. 5 Levels (Derived from Hisp_Cat and Race_Cat2)

The RaceEthn5 was created with Hisp_Cat and Race_Cat2 variables. The RaceEthn5 has 5 categories: Non-Hispanic White; Non-Hispanic Black or African American; Hispanic; Non-Hispanic Asian; Non-Hispanic Other. If Hisp_Cat had value of 10, "Not Hispanic", and Race_Cat2 had value of -9, "Missing data (Not Ascertained)", then RaceEthn5 was assigned with value of -9. If Hisp_Cat had value of 10, "Not Hispanic", and Race_Cat2 had value of -7, "Missing data (Web partial - Question Never Seen)", then RaceEthn5 was assigned with value of -7. RaceEthn5 was assigned with value of -9 if Hisp_Cat =-9. RaceEthn5 was assigned a value of -7 if Hisp_Cat=-7.

HHInc: What is your {combined} annual household income? 5 Levels (Derived from IncomeRanges)

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: Body Mass Index (Weight*703)/(Height in inches**2)

The BMI variable was created with weight in pounds and height in inches. If height in feet or weight had value of -9, "Missing data (Not Ascertained)" but neither had value of -4, "Unreadable or Non-conforming numeric response", then BMI was assigned to -9. If height in feet had a value greater than 0 and height in inches had a value of -4, "Unreadable or Non-conforming numeric response", then BMI was assigned to -9. If height in feet or weight had value of -4, "Unreadable or Non-conforming numeric

response", the BMI was assigned to -4. If height in feet or weight had value of -7, "Missing data (Web Partial- Question never seen)", then BMI was assigned to -7.

TimeSinceDX: How long ago were you diagnosed with cancer? (Derived from WhenDiagnosedCancer and Age)

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 a 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 a value of -7, "Missing data (Web partial- Question never seen)", TimeSinceDX was assigned to -7. 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. If the variable EverHadCancer had value of 1 and WhenDiagnosedCancer had value of -7, "Missing data (Web partial- Question never seen)", TimeSinceDX was assigned to -7.

SmokeStat: Smoking Status (Derived from Smoke100 and SmokeNow)

The variable smokeStat was created with Smoke100 and SmokeNow variables. The variable smokeStat has 3 categories: Current; Former; Never. Starting in HINTS 6, SmokeNow was administered regardless of the response to Smoke100. In prior HINTS surveys, SmokeNow was only administered if the response of Smoke100 was Yes. Due to this skip pattern difference, SmokeStat was derived differently than in previous versions. If Smoke100 had a value of 1 or 2 (Yes or No) and SmokeNow had a value of 1 or 2 (Every day or Some days), then SmokeStat was assigned a value of 1 (Current). If Smoke100 had a value of 1 (Yes) and SmokeNow had a value of 3 (Not at all), then SmokeStat was assigned a value of 2 (Former). If Smoke100 had a value of 2 (No) and SmokeNow had a value of 3 (Not at all), then SmokeStat was assigned a value of 3 (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 or 2 and SmokeNow had value of -9, "Missing data (Not Ascertained)", then smokeStat was assigned to -9. If Smoke100 had value of -9, "Missing data (Not Ascertained)", the smokeStat was assigned to -6. If Smoke100 had value of -7, "Missing data (Web partial - Question Never Seen)" and SmokeNow had value of -7, "Missing data (Web partial - Question Never Seen)" then smokeStat was assigned to -7.

PHQ4: PHQ-4 total score (Derived composite from LittleInterest, Hopeless, Nervous, and Worrying)

The variable PHQ4 was created with LittleInterest, Hopeless, Nervous and Worrying variables. We created total score as continuous variable: 1) Rescore variables 0-3 and then reverse coding such that 'Not at all=0', 'Several Days'=1, 'More than half the days'=2, 'Nearly every day'=3 2) Compute total score by summing across 4 items 3) Total score range will be 0-12. If one of LittleInterest, Hopeless, Nervous and Worrying variables had value of -5, "Multiple response selected in error", the PHQ was assigned to -5. If one of LittleInterest, Hopeless, Nervous and Worrying variables had a value of -9, "Missing data (Not Ascertained)", then PHQ was assigned to -9. If one of LittleInterest, Hopeless, Nervous and Worrying

variables had a value of -7, “Missing data (Web partial - Question Never Seen)”, then PHQ was assigned to -7.

WeeklyMinutesModerateExercise: Minutes per week of at least moderate intensity exercise (Derived from TimesModerateExercise and HowLongModerateExerciseMinutes)

The variable WeeklyMinutesModerateExercise was created with TimesModerateExercise and HowLongModerateExerciseMinutes variables. If TimesModerateExercise is 0 then WeeklyMinutesModerateExercise was assigned to 0. If TimesModerateExercise is less than 0 then was assigned to the value of TimesModerateExercise (i.e., the original negative values were retained). If HowLongModerateExerciseMinutes is less than 0 then WeeklyMinutesModerateExercise was assigned to the value of HowLongModerateExerciseMinutes (i.e., the original negative values were retained).

eCigUse: Electronic Cigarette Use (Derived from UsedECigEver and UseECigNow Recode)

The variable eCigUse was created with UsedECigEver and UseECigNow variables. The variable eCigUse has 3 categories: Current; Former; Never. If UsedECigEver had value of 1 and UseECigNow had value of -5, “Multiple responses selected in error”, the eCigUse was assigned to -4. If UsedECigEver had value of 1 and UseECigNow had value of -9, “Missing data (Not Ascertained)”, the eCigUse was assigned to -9. If UsedECigEver had value of -9, “Missing data (Not Ascertained)”, the eCigUse was assigned to -6. If UsedECigEver or UseECigNow had value of -7, “Missing data (Web Partial- Question Never Seen)”, then eCigUse was assigned to -7.

AvgDrinksPerWeek: Average number of drinks per week (Derived from DrinkDaysPerWeek and DrinksPerDay)

The variable AvgDrinksPerWeek was created with DrinkDaysPerWeek and DrinksPerDay variables. If DrinkDaysPerWeek is less than 0 then AvgDrinksPerWeek was assigned to the value of DrinkDaysPerWeek (i.e., the original negative values were retained). If DrinksPerDay is less than 0 then AvgDrinksPerWeek was assigned to the value of DrinksPerDay (i.e., the original negative values were retained).

PCCScale: Patient Centered Communication scale

The PCCScale variable was created with ChanceAskQuestions, FeelingsAddressed, InvolvedDecisions, UnderstoodNextsteps, ExplainedClearly, SpentEnoughTime, and HelpUncertainty. We reverse the values of these 7 variables and take the mean of these 7 variables (if at least half the variables have valid values). And then we transform the mean value linearly to a 0-100 scale. If at least half the input variables did not have valid values and any of the 7 input items were -9, then PCCScale was assigned a value of -9. If at least half the input variables did not have valid values and any of the 7 input items were -7, “Missing data (Web Partial- Question Never Seen)”, then PCCScale was assigned a value of -7. If at least half the input variables did not have valid values and the FreqGoProvider variable was 0, “None”, then PCCScale was assigned a value of -1. See for more information about the psychometric properties of the PCC scale: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9803193/

ISEE_Scale: Information Seeking Experience scale

The ISEE_Scale variable was created with CancerLotOfEffort, CancerFrustrated, CancerConcernedQuality and CancerTooHardUnderstand. We take the mean of these 4 variables (if all the variables have valid

values). And then we transform the mean value linearly to a 0-100 scale. If any of the 4 input items for ISEE_Scale were -9, then ISEE_Scale was assigned a value of -9. If any of the 4 input items for ISEE_Scale were -7, “Missing data (Web Partial- Question Never Seen)”, then ISEE_Scale was assigned a value of -7. If the SeekCancerInfo variable was 2, “No” then ISEE_Scale was assigned a value of -1.

PROMIS_Meaning_t: PROMIS Meaning and Purpose in Life T scores

The PROMIS_Meaning_t variable was created with LifeHasMeaning, ClearSenseDir, DeepFulfillment and LifeHasPurpose. We reverse score the values of these 4 variables and take the sum of these 4 variables (if all variables have valid values) and then we transform the sum score to a T-score (which has a mean of 50 and standard deviation of 10) using the PROMIS scoring conversion table. If any of the 4 input variables for PROMIS_Meaning_t were -9, then PROMIS_Meaning_t was assigned a value of -9. If any of the 4 input variables for PROMIS_Meaning_t were -7, “Missing data (Web Partial- Question Never Seen)”, then PROMIS_Meaning_t was assigned a value of -7. See for more information regarding PROMIS measures: https://www.healthmeasures.net/explore-measurement-systems/promis

PROMIS_Isolation_t: PROMIS Social Isolation scale T scores

The PROMIS_Isolation_t variable was created with FeelLeftOut, FeelPeopleBarelyKnow, FeelIsolated and FeelPeopleNotWithMe. We take the sum of these 4 variables (if all variables have valid values) and then we transform the sum score to T-score (which has a mean of 50 and standard deviation of 10) using the PROMIS scoring conversion table. If any of the 4 input variables for PROMIS_Isolation_t were -9, then PROMIS_Isolation_t was assigned a value of -9. If any of the 4 input variables for PROMIS_Isolation_t were -7, “Missing data (Web Partial- Question Never Seen)”, then PROMIS_Isolation_t was assigned a value of -7. See for more information regarding PROMIS measures: https://www.healthmeasures.net/explore-measurement-systems/promis

Label Editing

Labels Added for Standard Recode Variables

Labels were created for the following recoded variables:, AgeGrpA, AgeGrpB, EducA, EducB, RaceEthn, RaceEthn5, HHInc, BMI, TimeSinceDX, smokeStat, PHQ4, WeeklyMinutesModerateExercise, eCigUse, AvgDrinksPerWeek, PCCScale, ISEE_Scale, PROMIS_Meaning_t, and PROMIS_Isolation_t.

Labels were created for the following variables that did not originally have labels on the dataset: EverHadCancer and PROMPT.

Labels Modified for Certain Variables

Labels were modified for the following variables to correct typos and improve clarity:HadTest3_NotHad, THNo_ConcernedPrivacy, THNo_TooDifficult, ClearSenseDir, DeepFulfillment, LifeHasPurpose, and LackTransportation.

Format Editing

Formats Added for Standard Recode Variables

The formats AgeGrpA, AgeGrpB, EducA, EducB, RaceEthn, RaceE5f, HHInc, BMI, TSDX, sStat, phq4f, WMWE, ecigS, ADPW, pccscale, ISEscale, PROMISm, and PROMISi were created and assigned to the variables AgeGrpA, AgeGrpB, EducA, EducB, RaceEthn, RaceEthn5, HHInc, BMI, TimeSinceDX, smokeStat, PHQ4, WeeklyMinutesModerateExercise, eCigUse, AvgDrinksPerWeek, PCCScale, ISEE_Scale, PROMIS_Meaning_t, and PROMIS_Isolation_t respectively.

Formats Modified for Certain Variables

All skip patterns in formats were modified (i.e. all instructions to skip questions were deleted). The modified formats were: Adultsf, SeekCf, UseIntf, HD_Cellf, HD_Nonef, HD_Catf, WDTHeaf, MlHf, FGPf, RTCf, OTOf, AOR2f, MOPsf, C_Nof, HGT_Nf, TS_NHf, HT3_NSf, HT3_NHf, GTEf, TSunf, DDPWf, HCPAC2f, TMEf, UECigEf, HHPVf, EHCf.

All "&" symbols in formats were removed. The modified formats were: DMAf.

Formats Removed for Dropped Variables

The following formats were removed to reflect removal of variables: HIGHSPf for removed variable HIGHSPANLI, HISPSURf for removed varable HISPSURNAME, HISP_HHf for removed variable HISP_HH, Qdispf for removed variable QDisp, and $PerIDf for removed variable PersonID.

Imputation of Income Variable

The income variable (IncomeRanges) has relatively higher percentage (12% for un-weighted percentage or 11% for weighted percentage) of missing values. This variable was imputed via PROC IMPUTE in SUDAAN. The imputation class variables are: Education (R7), RaceEthn (standard recode), and FullTimeOcc_Cat (derived variable from R4 and R5). 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=0 then MailHHAdults = 1;

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

* Recode HowLongExerciseMinutes 999 to missing;
if HowLongModerateExerciseMinutes=999 then HowLongModerateExerciseMinutes=-4;

* Recode Totalhousehold of 0 to 1;
if Totalhousehold=0 then Totalhousehold=1;

* Recode outliers of AverageTimeSitting (>20 hours)*;
if AverageTimeSitting > 20 then AverageTimeSitting = -4;
```

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 (-9,-4) then AgeGrpA=Age;
label AgeGrpA='AgeGrpA: Respondent Age Recode-4 Levels (Derived from
Age; see History Document for more information)';


if 18<=Age<=34 then AgeGrpB=1;
if 35<=Age<=49 then AgeGrpB=2;
if 50<=Age<=64 then AgeGrpB=3;
if 65<=Age<=74 then AgeGrpB=4;
if 75<=Age then AgeGrpB=5;
else if Age in (-9,-4) then AgeGrpB=Age;
label AgeGrpB='AgeGrpB: Respondent Age Recode-5 Levels (Derived from
Age; see History Document for more information)';


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, -7) then
      EducA = Education;
label EducA = 'EducA:  What is the highest level of school you
completed? (Education recoded-4 levels) (Derived from Education; see
History Document for more information)';


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, -7) then
      EducB = Education;
label EducB = 'EducB:  What is the highest level of school you
completed? (Education recoded-5 levels) (Derived from Education; see
History Document for more information)';


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= -7 then
                  RaceEthn = -7;
            else if Race_Cat2= -9 then
                  RaceEthn = -9;
            end;
else if Hisp_Cat= -7 then
      RaceEthn = -7;
else if Hisp_Cat= -9 then
            RaceEthn = -9;
label RaceEthn = 'Race/Ethnicity recode (Hisp_cat and Race_cat2--7
levels) (Derived from Hisp_Cat and Race_Cat2; see History Document for
more information)';


if Hisp_Cat in (21, 22, 23, 24, 25) then RaceEthn5 = 3;
else if Hisp_Cat in (10) then do;
if Race_Cat2 in (11) then RaceEthn5 = 1;
else if Race_Cat2 in (12) then RaceEthn5 = 2;
else if Race_Cat2 in (31, 32, 33, 34, 35, 36, 37) then RaceEthn5 = 4;
else if Race_Cat2 in (51, 52, 53, 54,14,16) then RaceEthn5 = 5;
else if Race_Cat2= (-7) then RaceEthn5 = -7;
else if Race_Cat2= (-9) then RaceEthn5 = -9;
end;
if Hisp_cat= -7 then RaceEthn5=-7;
if Hisp_Cat = -9 then RaceEthn5 = -9;
label RaceEthn5 = 'Race/Ethnicity recode (Hisp_cat and Race_cat2--5
levels) (Derived from Hisp_Cat and Race_Cat2; see History Document for
more information)';


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, -7) then
      HHInc = IncomeRanges;
label HHInc = 'HHInc:  What is your (combined) annual household income?
(IncomeRanges Recode-5 levels)(Derived from IncomeRanges; see History
Document for more information)';


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 in (-9) and Weight >=-1 ) or (Height_Feet >=-1 and
Weight in (-9)) or (Height_Feet = -9 and Weight = -9) or (Height_Feet >
0 and Height_Inches = -4) then
      BMI = -9;
else if Height_Feet = -4 or Weight = -4 then
      BMI = -4;
else if Height_Feet = -7 or Weight = -7 then
      BMI = -7;
label BMI = 'BMI.  Body Mass Index (Weight*703)/(Height in inches**2)
(See History Document for more information)';
if BMI not in (-4, -7, -9) then
      BMI = round(BMI, 0.1);


if EverHadCancer in (1) then do;
      if Age < 0 then
            TimeSinceDX = Age;
      else if WhenDiagnosedCancer <0 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 (-7) then
      TimeSinceDX = WhenDiagnosedCancer;
else if EverHadCancer in (-9) then
      TimeSinceDX = WhenDiagnosedCancer;
else if EverHadCancer in (2) then do;
            if WhenDiagnosedCancer in (-1) then
                  TimeSinceDX = WhenDiagnosedCancer;
            if WhenDiagnosedCancer in (-7) then
                  TimeSinceDX = WhenDiagnosedCancer;
            else if WhenDiagnosedCancer in (-2) then
                  TimeSinceDX = -4;
end;
label TimeSinceDX = 'TimeSinceDX: How long ago were you diagnosed with
cancer? (Derived from WhenDiagnosedCancer and Age; see History Document
for more information)';


/*New skip logic in Smoke100/SmokeNow for HINTS 6*/
if (Smoke100 in (1) and  SmokeNow in (1,2)) or (Smoke100 in (2) and
SmokeNow in (1,2)) then smokeStat = 1;
else if Smoke100 in (1) and SmokeNow in (3) then smokeStat = 2;
else if Smoke100 in (2) and SmokeNow in (3) then  smokeStat = 3;
else if Smoke100 in (-9) then smokeStat = -6;
```

```
else if SmokeNow in (-5) then smokeStat = -4;
else if SmokeNow in (-7) or Smoke100 in (-7) then smokeStat = -7;
else if SmokeNow in (-9) or Smoke100 in (-9) then smokeStat = -9;
label smokeStat = 'SmokeStat: Smoking Status (Derived from Smoke100 and
SmokeNow; see History Document for more information)';


array b(*) LittleInterest Hopeless Nervous Worrying;
PHQ4= 0;
if b(1) = -5 or b(2) = -5 or b(3) = -5 or b(4) = -5 then PHQ4 =-5;
else if b(1) in (-7) or b(2) in (-7) or b(3) in (-7) or b(4) in  (-7)
then PHQ4 = -7;
else if b(1) in (-9) or b(2) in (-9) or b(3) in (-9) or b(4) in  (-9)
then PHQ4 = -9;
else do;
      do I = 1 to dim(b);
            if PHQ4 not in (-5, -7, -9) and b(I) in (1, 2,3 ,4) then
            PHQ4 = PHQ4 + (4-b(I));
      end;
end;
label PHQ4 = 'PHQ4: PHQ-4 total score (Derived composite from
LittleInterest, Hopeless, Nervous, and Worrying; see History Document
for more information)';
drop I;


If TimesModerateExercise=0 then WeeklyMinutesModerateExercise=0;
else If TimesModerateExercise<0 then
WeeklyMinutesModerateExercise=TimesModerateExercise;
else If HowLongModerateExerciseMinutes<0 then
WeeklyMinutesModerateExercise=HowLongModerateExerciseMinutes;
else IF TimesModerateExercise>0 then do;
      if HowLongModerateExerciseMinutes = 0 then
WeeklyMinutesModerateExercise = 0;
      else WeeklyMinutesModerateExercise =
HowLongModerateExerciseMinutes*TimesModerateExercise;
end;
label WeeklyMinutesModerateExercise="WeeklyMinutesModerateExercise:
Minutes per week of at least moderate intensity exercise (Derived from
TimesModerateExercise and HowLongModerateExerciseMinutes; see History
Document for more information)";


if UsedECigEver in (1) then
      do;
            if UseECigNow in (1, 2) then
                  eCigUse = 1;
            else if UseECigNow = 3 then
                  eCigUse = 2;
            else if UseECigNow = (-5) then
                  eCigUse = -4;
            else if UseECigNow in (-7) then
                  eCigUse = -7;
            else if UseECigNow in (-9) then
                  eCigUse = -9;
      end;
else if UsedECigEver in (2) then
```

```
            eCigUse = 3;
else if UsedECigEver in (-7) then
      eCigUse = -7;
else if UsedECigEver in (-9) then
      eCigUse = -6;
label eCigUse = 'eCigUse:  Electronic Cigarette Use (Derived from
UsedECigEver and UseECigNow; see History Document for more
information)';


if DrinkDaysPerWeek=0 then AvgDrinksPerWeek=0;
else if DrinkDaysPerWeek < 0 then AvgDrinksPerWeek=DrinkDaysPerWeek;
else if DrinksPerDay < 0 then AvgDrinksPerWeek=DrinksPerDay;
else if DrinkDaysPerWeek>0 then do;
      if DrinksPerDay=0 then AvgDrinksPerWeek = 0;
      else AvgDrinksPerWeek = DrinkDaysPerWeek*DrinksPerDay;
end;
if AvgDrinksPerWeek >= 140 then AvgDrinksPerWeek=-4;
label AvgDrinksPerWeek = "AvgDrinksPerWeek: Average number of drinks
per week (Derived from DrinkDaysPerWeek and DrinksPerDay; see History
Document for more information)";

/*Creating PCC Scale*/
if ChanceAskQuestions < 0 then ChanceAskQuestionsRev=.;
if ChanceAskQuestions = 1 then ChanceAskQuestionsRev=4;
if ChanceAskQuestions = 2 then ChanceAskQuestionsRev=3;
if ChanceAskQuestions = 3 then ChanceAskQuestionsRev=2;
if ChanceAskQuestions = 4 then ChanceAskQuestionsRev=1;
**************************;
if FeelingsAddressed < 0 then FeelingsAddressedRev=.;
if FeelingsAddressed = 1 then FeelingsAddressedRev=4;
if FeelingsAddressed = 2 then FeelingsAddressedRev=3;
if FeelingsAddressed = 3 then FeelingsAddressedRev=2;
if FeelingsAddressed = 4 then FeelingsAddressedRev=1;
**************************;
if InvolvedDecisions < 0 then InvolvedDecisionsRev=.;
if InvolvedDecisions = 1 then InvolvedDecisionsRev=4;
if InvolvedDecisions = 2 then InvolvedDecisionsRev=3;
if InvolvedDecisions = 3 then InvolvedDecisionsRev=2;
if InvolvedDecisions = 4 then InvolvedDecisionsRev=1;
**************************;
if UnderstoodNextSteps < 0 then UnderstoodNextStepsRev=.;
if UnderstoodNextSteps = 1 then UnderstoodNextStepsRev=4;
if UnderstoodNextSteps = 2 then UnderstoodNextStepsRev=3;
if UnderstoodNextSteps = 3 then UnderstoodNextStepsRev=2;
if UnderstoodNextSteps = 4 then UnderstoodNextStepsRev=1;
**************************;
if ExplainedClearly < 0 then ExplainedClearlyRev=.;
if ExplainedClearly = 1 then ExplainedClearlyRev=4;
if ExplainedClearly = 2 then ExplainedClearlyRev=3;
if ExplainedClearly = 3 then ExplainedClearlyRev=2;
if ExplainedClearly = 4 then ExplainedClearlyRev=1;
**************************;
if SpentEnoughTime < 0 then SpentEnoughTimeRev=.;
if SpentEnoughTime = 1 then SpentEnoughTimeRev=4;
if SpentEnoughTime = 2 then SpentEnoughTimeRev=3;
if SpentEnoughTime = 3 then SpentEnoughTimeRev=2;
```

```
if SpentEnoughTime = 4 then SpentEnoughTimeRev=1;
**************************;
if HelpUncertainty < 0 then HelpUncertaintyRev=.;
if HelpUncertainty = 1 then HelpUncertaintyRev=4;
if HelpUncertainty = 2 then HelpUncertaintyRev=3;
if HelpUncertainty = 3 then HelpUncertaintyRev=2;
if HelpUncertainty = 4 then HelpUncertaintyRev=1;
**************************;

** Create the PCCMean variable;
If nmiss
(ChanceAskQuestionsRev,FeelingsAddressedRev,InvolvedDecisionsRev,
UnderstoodNextStepsRev,ExplainedClearlyRev,SpentEnoughTimeRev,
HelpUncertaintyRev)<4
then PCCMean =
MEAN(ChanceAskQuestionsRev,FeelingsAddressedRev,InvolvedDecisionsRev,
UnderstoodNextStepsRev,ExplainedClearlyRev,SpentEnoughTimeRev,
HelpUncertaintyRev);

/*Create composite score - linear variable on a scale of 100*/
PCCScale=round(((pccmean-1)*100)/3,0.1);
if PCCScale < 0 then PCCScale=.;
if ChanceAskQuestions=-7 or FeelingsAddressed=-7 or InvolvedDecisions=-
7 or UnderstoodNextSteps=-7 or ExplainedClearly=-7 or SpentEnoughTime=-
7 or HelpUncertainty=-7 then PCCScale=-7;
if PCCScale=. then PCCScale=-9;
if FreqGoProvider=0 then PCCScale=-1;
label PCCScale='PCCScale: Patient Centered Communication scale';

/*Create ISEE_scale */
** create the ISEEMean variable;
If nmiss (CancerLotOfEffort,CancerFrustrated,CancerConcernedQuality,
CancerTooHardUnderstand)<1
      Then ISEE_Mean =
MEAN(CancerLotOfEffort,CancerFrustrated,CancerConcernedQuality,
CancerTooHardUnderstand);

/*Create composite score - linear variable on a scale of 100*/
ISEE_Scale=round(((ISEE_mean-1)*100)/3,0.1);

if ISEE_Scale < 0 then ISEE_Scale=.;
if CancerLotOfEffort=-7 or CancerFrustrated=-7 or
CancerConcernedQuality=-7 or CancerTooHardUnderstand=-7 then
ISEE_Scale=-7;
if ISEE_Scale=. then ISEE_Scale=-9;
if SeekCancerInfo=2 then ISEE_Scale=-1;
label ISEE_Scale='ISEE_Scale: Information Seeking Experience scale';


/*PROMIS Social Isolation*/
if feelleftout<0 or feelpeoplebarelyknow<0 or feelisolated<0 or
feelpeoplenotwithme<0 then PROMIS_Isolation_t=.;
else do;
      SIsum=sum(feelleftout,feelpeoplebarelyknow,feelisolated,
      feelpeoplenotwithme);
           if SIsum=4 then PROMIS_Isolation_t=34.8;
      else if SIsum=5 then PROMIS_Isolation_t=40.4;
```

```
      else if SIsum=6 then PROMIS_Isolation_t=43.3;
      else if SIsum=7 then PROMIS_Isolation_t=45.7;
      else if SIsum=8 then PROMIS_Isolation_t=47.8;
      else if SIsum=9 then PROMIS_Isolation_t=49.8;
      else if SIsum=10 then PROMIS_Isolation_t=51.8;
      else if SIsum=11 then PROMIS_Isolation_t=53.9;
      else if SIsum=12 then PROMIS_Isolation_t=56.1;
      else if SIsum=13 then PROMIS_Isolation_t=58.1;
      else if SIsum=14 then PROMIS_Isolation_t=60.1;
      else if SIsum=15 then PROMIS_Isolation_t=62.0;
      else if SIsum=16 then PROMIS_Isolation_t=63.8;
      else if SIsum=17 then PROMIS_Isolation_t=65.5;
      else if SIsum=18 then PROMIS_Isolation_t=67.5;
      else if SIsum=19 then PROMIS_Isolation_t=69.9;
      else if SIsum=20 then PROMIS_Isolation_t=74.2;
end;
if feelleftout=-7 or feelpeoplebarelyknow=-7 or feelisolated=-7 or
feelpeoplenotwithme=-7 then PROMIS_Isolation_t=-7;
if PROMIS_Isolation_t=. then PROMIS_Isolation_t=-9;
label PROMIS_Isolation_t="PROMIS Social Isolation Scale T Scores
(Derived from FeelLeftOut, FeelPeopleBarelyKnow, FeelIsolated,
FeelPeopleNotWithMe; see History Document for more information)";

/* PROMIS Meaning */
if LifeHasMeaning < 0 then LifeHasMeaningRev=.;
if LifeHasMeaning = 1 then LifeHasMeaningRev=5;
if LifeHasMeaning = 2 then LifeHasMeaningRev=4;
if LifeHasMeaning = 3 then LifeHasMeaningRev=3;
if LifeHasMeaning = 4 then LifeHasMeaningRev=2;
if LifeHasMeaning = 5 then LifeHasMeaningRev=1;
**************************;
if ClearSenseDir < 0 then ClearSenseDirRev=.;
if ClearSenseDir = 1 then ClearSenseDirRev=5;
if ClearSenseDir = 2 then ClearSenseDirRev=4;
if ClearSenseDir = 3 then ClearSenseDirRev=3;
if ClearSenseDir = 4 then ClearSenseDirRev=2;
if ClearSenseDir = 5 then ClearSenseDirRev=1;
**************************;
if DeepFulfillment < 0 then DeepFulfillmentRev=.;
if DeepFulfillment = 1 then DeepFulfillmentRev=5;
if DeepFulfillment = 2 then DeepFulfillmentRev=4;
if DeepFulfillment = 3 then DeepFulfillmentRev=3;
if DeepFulfillment = 4 then DeepFulfillmentRev=2;
if DeepFulfillment = 5 then DeepFulfillmentRev=1;
**************************;
if LifeHasPurpose < 0 then LifeHasPurposeRev=.;
if LifeHasPurpose = 1 then LifeHasPurposeRev=5;
if LifeHasPurpose = 2 then LifeHasPurposeRev=4;
if LifeHasPurpose = 3 then LifeHasPurposeRev=3;
if LifeHasPurpose = 4 then LifeHasPurposeRev=2;
if LifeHasPurpose = 5 then LifeHasPurposeRev=1;
**************************;

if LifeHasMeaning<0 or ClearSenseDir<0 or DeepFulfillment<0 or
LifeHasPurpose<0 then PROMIS_Meaning_t=.;
else do;
```

```
      PMsum=sum(LifeHasMeaningRev,ClearSenseDirRev,DeepFulfillmentRev,L
ifeHasPurposeRev);
             if PMsum=4 then PROMIS_Meaning_t=21.2;
      else if PMsum=5 then PROMIS_Meaning_t =25.2;
      else if PMsum=6 then PROMIS_Meaning_t =28.2;
      else if PMsum=7 then PROMIS_Meaning_t =30.7;
      else if PMsum=8 then PROMIS_Meaning_t =33.0;
      else if PMsum=9 then PROMIS_Meaning_t =35.2;
      else if PMsum=10 then PROMIS_Meaning_t =37.4;
      else if PMsum=11 then PROMIS_Meaning_t =39.5;
      else if PMsum=12 then PROMIS_Meaning_t =41.7;
      else if PMsum=13 then PROMIS_Meaning_t =43.9;
      else if PMsum=14 then PROMIS_Meaning_t =46.1;
      else if PMsum=15 then PROMIS_Meaning_t =48.4;
      else if PMsum=16 then PROMIS_Meaning_t =50.8;
      else if PMsum=17 then PROMIS_Meaning_t =53.5;
      else if PMsum=18 then PROMIS_Meaning_t =56.5;
      else if PMsum=19 then PROMIS_Meaning_t =60.0;
      else if PMsum=20 then PROMIS_Meaning_t =65.5;
end;
if LifeHasMeaning=-7 or ClearSenseDir=-7 or DeepFulfillment=-7 or
LifeHasPurpose=-7 then PROMIS_Meaning_t=-7;
if PROMIS_Meaning_t=. then PROMIS_Meaning_t=-9;
label PROMIS_Meaning_t ="PROMIS Meaning and Purpose in Life T Scores
(Derived from LifeHasMeaning, ClearSenseDir, DeepFulfillment,
LifeHasPurpose; see History Document for more information)";
```

SAS Code for Format Editing

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"
```

```
-7 = "Missing data (Web partial - Question Never Seen)"
-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"
-7 = "Missing data (Web partial - Question Never Seen)"
-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'
-7 = 'Missing data (Web partial - Question Never Seen)'
-9 = 'Missing Data (Not Ascertained)'
;

value RaceE5f
1 = "Non-Hispanic White"
2 = "Non-Hispanic Black or African American"
3 = "Hispanic"
4 = "Non-Hispanic Asian"
5 = "Non-Hispanic Other"
-7 = "Missing data (Web partial - Question Never Seen)"
-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"
-7 = "Missing data (Web partial - Question Never Seen)"
-9 = "Missing Data (Not Ascertained) "
;

value BMI
-4 = 'Unreadable or Nonconforming Numeric Response'
-7 = 'Missing data (Web partial - Question Never Seen)'
-9 = 'Missing Data (Not Ascertained)'
;

value TSDX
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'
-7 = 'Missing data (Web partial - Question Never Seen)'
-9 = 'Missing Data (Not Ascertained)'
;

value sStat
1 = 'Current'
2 = 'Former'
3 = 'Never'
-4 = 'Unreadable or Nonconforming Numeric Response'
-6 = 'Missing Data (Filter Missing), coded -9 in Smoke100'
-7 = 'Missing data (Web partial - Question Never Seen)'
-9 = 'Missing Data (Not Ascertained)'
;

value phq4f
-5 = 'Multiple Responses Selected in Error'
-7 = 'Missing data (Web partial - Question Never Seen)'
-9 = 'Missing Data (Not Ascertained)'
;

value WMWE
-4 = 'Unreadable or Nonconforming Numeric Response'
-5 = 'Multiple Responses Selected in Error'
-7 = 'Missing data (Web partial - Question Never Seen)'
-9 = 'Missing Data (Not Ascertained)'
;

value ecigS
1 = 'Current'
2 = 'Former'
3 = 'Never'
-4 = 'Unreadable or Nonconforming Numeric Response'
-6 = 'Missing Data (Filter Missing), coded -9 in UsedECigEver'
-7 = 'Missing data (Web partial - Question Never Seen)'
-9 = 'Missing Data (Not Ascertained)'
;

value ADPW
-4 = 'Unreadable or Nonconforming Numeric Response'
-5 = 'Multiple Responses Selected in Error'
-7 = 'Missing data (Web partial - Question Never Seen)'
-9 = 'Missing Data (Not Ascertained)'
;

value pccscale
-1 = 'Inapplicable, coded 0 in FreqGoProvider'
-7 = 'Missing data (Web partial - Question Never Seen)'
-9 = 'Missing data (Not Ascertained)'
;

value ISEscale
-1 = 'Inapplicable, coded 2 in SeekCancerInfo'
```

```
-7 = 'Missing data (Web partial - Question Never Seen)'
-9 = 'Missing data (Not Ascertained)'
;

value PROMISm
-7 = 'Missing data (Web partial - Question Never Seen)'
-9 = 'Missing data (Not Ascertained)'
;

value PROMISi
-7 = 'Missing data (Web partial - Question Never Seen)'
-9 = 'Missing data (Not Ascertained)'
;

format      AgeGrpA AgeGrpA.
            AgeGrpB AgeGrpB.
            EducA EducA.
            EducB EducB.
            RaceEthn RaceEthn.
            RaceEthn5 RaceE5f.
            HHInc HHInc.
            BMI BMI.
            TimeSinceDX TSDX.
            smokeStat sStat.
            phq4 phq4f.
            WeeklyMinutesModerateExercise WMWE.
            ECigUse ecigS.
            IncomeRanges_IMP IRf.
            AvgDrinksPerWeek ADPW.
            PCCScale pccscale.
            ISEE_Scale ISEscale.
            PROMIS_Meaning_t PROMISm.
            PROMIS_Isolation_t PROMISi.
;
```

SAS Code for Imputation of Income Variable

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

      COPY_Education = Education;
      if COPY_Education in (-9, -7) then
            COPY_Education = .;

      COPY_RaceEthn = RaceEthn;
      if COPY_RaceEthn in (-9, -7) then
            COPY_RaceEthn = .;

/*    COPY_RentOrOwn = RentOrOwn;
      if COPY_RentOrOwn in (-5, -7, -9) then
            COPY_RentOrOwn = .;

      COPY_SpeakEnglish = SpeakEnglish;
      if COPY_SpeakEnglish in (-5, -7, -9) then
            COPY_SpeakEnglish = .;*/
```

```
      COPY_IncomeRanges = IncomeRanges;
      if COPY_IncomeRanges in (-7, -9) then
            COPY_IncomeRanges = .;

    COPY_FullTimeOcc_Cat = FullTimeOcc_Cat;
      if FullTimeOcc_Cat in (-4, -9, -7) then
            COPY_FullTimeOcc_Cat = .;

    ID = _N_;

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

proc freq data=hints;
      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_IncomeRanges*IncomeRanges / list missing;
      tables COPY_FullTimeOcc_Cat*FullTimeOcc_Cat/ list missing;
run;

proc impute data=hints method=wshd notsorted;
      weight PERSON_FINWT0;
      impvar COPY_IncomeRanges;
      impby COPY_Education COPY_RaceEthn COPY_FullTimeOcc_Cat;
      impname COPY_IncomeRanges="IncomeRanges_IMP";
      impid ID;
      output IMPID IMPBY IMPUTEVAL / filename=impute1 replace;
run;

proc freq data=impute1;
      tables IncomeRanges_IMP / missing;
run;

proc contents data=impute1;
run;

proc sort data=hints;
      by ID;
run;

proc sort data=impute1 (keep=ID IncomeRanges_IMP);
      by ID;
run;

data hints;
      merge hints (in=A) impute1 (in=B);
      by ID;

      if A = 1 and B = 1;
run;

data _null_;
      set hints;
```

```
      if IncomeRanges not in (-7 , -9) and COPY_IncomeRanges ^=
      IncomeRanges_IMP then
            put ID IncomeRanges COPY_IncomeRanges IncomeRanges_IMP;
run;

data hints;
      set hints;

      if missing(IncomeRanges_IMP) = 1 then
            IncomeRanges_IMP = IncomeRanges;
      label IncomeRanges_IMP = '-->IncomeRanges_IMP.  Imputed IncomeRanges
variable via PROC HOTDECK in SUDAAN (see History Document for more
information)';

      drop COPY_Education COPY_RaceEthn /*COPY_RentOrOwn COPY_SpeakEnglish*/
      COPY_FullTimeOcc_Cat
      ID
      COPY_IncomeRanges;
run;
```