SmokingPackYears_MP

*--------------------------------*  
| PROGRAM NAME:               | SMOKINGPACKYEARS_MP.SAS     |  
| VARIABLES CREATED:         | PACKYEARS: PY0 PY2 PY5 PY7 PY10 PY15 |  
| LTSMK5 LTSMK7              | PY20 PY25 LTSMK0 LTSMK2        |  
| LTSMK25 LTSMK30            | LTSMK10 LTSMK15 LTSMK20       |  
| CARDIA DATASETS USED:      | A5REF B4REF C2REF D3REF E3REF F3REF |  
|                           | G6REF H5REF I0REF             |  
|                           | A4F09TOB B2F09TOB C1F09TOB D1F9TOBA |  
|                           | E1F09TOB F1F09TOB G3F09TOB    |  
| H3F09TOB                  | I1F09TOB                      |  
|                           | A4F10 B2F10 C1F10 D1F10 E1F10 |  
|                           | F1F10 G3F10 H3F10 I1F10      |  
| VARIABLES USED:           | RACE SEX EXAMAGE EX2_AGE EX3_AGE |  
|                           | EX4_AGE EX5_AGE EX6_AGE      |  
| EX7_AGE EX8_AGE           | EX9_AGE                       |  
|                           | A09SMKyr A09CGTDY B09SMKNW B09STS3M |  
|                           | B09LS3M B09CGTDY C09SMKNW    |  
|                           | C09LS3M C09CGTDY D09SMKNW    |  
|                           | D09LS3M D09CGTDY E09SMKNW    |  
|                           | E09LS3M E09CGTDY F09SMKNW    |  
|                           | F09LS3M F09CGTDY G09SMKNW    |  
|                           | G09LS3M G09CGTDY H09SMKNW    |  
|                           | H09LS3M H09CGTDY I09ST3M     |  
|                           | I09CGTDY I09SMKNW A10CIGS A10TOBAC |  
|                           | A10CIGAR A10PIPE A10SMKLS B10CIGS |  
|                           | B10TOBAC B10CIGAR B10PIPE    |  
|                           | B10SMKLS                      |  
|                           | C10CIGS C10TOBAC C10CIGAR    |  
|                           | C10SMKLS D10CIGS D10TOBAC    |  
|                           | D10PIPE D10SMKLS E10CIGS     |  
|                           | A09SMKyr A09CGTDY B09SMKNW B09STS3M |  
|                           | B09LS3M B09CGTDY C09SMKNW    |  
|                           | C09LS3M C09CGTDY D09SMKNW    |  
|                           | D09LS3M D09CGTDY E09SMKNW    |  
|                           | E09LS3M E09CGTDY F09SMKNW    |  
|                           | F09LS3M F09CGTDY G09SMKNW    |  
|                           | G09LS3M G09CGTDY H09SMKNW    |  
|                           | H09LS3M H09CGTDY I09ST3M     |  
|                           | I09CGTDY I09SMKNW A10CIGS A10TOBAC |  
|                           | A10CIGAR A10PIPE A10SMKLS B10CIGS |  
|                           | B10TOBAC B10CIGAR B10PIPE    |  
|                           | B10SMKLS                      |  
|                           | C10CIGS C10TOBAC C10CIGAR    |  
|                           | C10SMKLS D10CIGS D10TOBAC    |  
|                           | D10PIPE D10SMKLS E10CIGS     |  
|                           | A09SMKyr A09CGTDY B09SMKNW B09STS3M |  
|                           | B09LS3M B09CGTDY C09SMKNW    |  
|                           | C09LS3M C09CGTDY D09SMKNW    |  
|                           | D09LS3M D09CGTDY E09SMKNW    |  
|                           | E09LS3M E09CGTDY F09SMKNW    |  
|                           | F09LS3M F09CGTDY G09SMKNW    |  
|                           | G09LS3M G09CGTDY H09SMKNW    |  
|                           | H09LS3M H09CGTDY I09ST3M     |  
|                           | I09CGTDY I09SMKNW A10CIGS A10TOBAC |  
|                           | A10CIGAR A10PIPE A10SMKLS B10CIGS |  
|                           | B10TOBAC B10CIGAR B10PIPE    |  
|                           | B10SMKLS                      |  
|                           | C10CIGS C10TOBAC C10CIGAR    |  
|                           | C10SMKLS D10CIGS D10TOBAC    |  
|                           | D10PIPE D10SMKLS E10CIGS     |  

Page 1
* SMOKING PACK YEARS BY DR. MARK PLETCHER;
* MAKE SMOKE DATA FOR DR. MARK PLETCHER;

* ALGORITHM:
* 1. SET 0 FOR SKIP-PATTERN;
* 2. FOR THOSE ANSWER "NO" TO CURRENT SMOKING, CODE CGTDY = 0;
* 3. FOR MISSING CIGS/DAY AT THE FIRST EXAM, ASSUME IT IS THE SAME AS THE FIRST NON-MISSING VALUE;
* 4. FOR MISSING DURATION OF SMOKING AT THE FIRST EXAM;
  * A) IF FIRST NON-MISSING CIGS/DAY IS 0, THEN ASSUME 0 (AND DOESN'T MATTER);
  * B) IF FIRST NON-MISSING CIGS/DAY IS NON-ZERO, THEN CALCULATE AVERAGE AGE OF SMOKING ONSET IN CARDIA FOR BASELINE SMOKERS;
  * C) IF AGE IS 17, 18, OR 19, SET A09SMKYR = 1;
* 5. FOR MISSING VALUES AT "MIDDLE" EXAMS (NOT FIRST OR LAST), ASSUME CONSTANT CHANGE IN PACK-YEARS
  * TO GET FROM PRE-MISSING TO POST-MISSING VALUES, INCORPORATE A SMOOTHER IMPUTATION FOR MISSING VALUES;
* 6. FOR MISSING VALUES AT THE "END" - ASSUME NO CHANGE IN CIGS/DAY SINCE LAST MEASUREMENT;

* BY FENG LIN
* ON 12/08/06;

************************************************************************************************************************************************************************
Page 2
DATA AGE;
  MERGE CARDIA.A5REF (KEEP=ID EXAMAGE)
    CARDIA.B4REF (KEEP=ID EX2_AGE)
    CARDIA.C2REF (KEEP=ID EX3_AGE)
    CARDIA.D3REF (KEEP=ID EX4_AGE)
    CARDIA.E3REF (KEEP=ID EX5_AGE)
    CARDIA.F3REF (KEEP=ID EX6_AGE)
    CARDIA.G6REF (KEEP=ID EX7_AGE)
    CARDIA.H5REF (KEEP=ID EX8_AGE)
    CARDIA.I1REF (KEEP=ID EX9_AGE);
  BY ID;
  * LAST VISIT WITH NONMISSING AGE ;
  ARRAY AGES{9} EXAMAGE EX2_AGE EX3_AGE EX4_AGE EX5_AGE EX6_AGE EX7_AGE EX8_AGE EX9_AGE;
  LASTVIS=0;
  DO I=9 TO 1 BY -1 WHILE (LASTVIS=0);
    IF AGES{I}>. THEN
      LASTVIS=I;
    END;
  KEEP ID LASTVIS;
  PROC SORT;
  BY ID;
RUN;

DATA F3REF;
  SET CARDIA.F3REF (KEEP=ID RACE SEX EXAMAGE);
RUN;

DATA ID;
  SET CARDIA.F3REF (KEEP=ID);
  SHORT_ID=INT(ID/10**7);
RUN;

DATA B2F09TOB (DROP=SHORT_ID);
  MERGE ID CARDIA.B2F09TOB (KEEP=SHORT_ID B09SMKNW B09STS3M B09SLS3M B09CGTDY);
  BY SHORT_ID;
RUN;

DATA TEMP;
  MERGE CARDIA.A4F09TOB (KEEP=ID A09SMKYR A09CGTDY)
SmokingPackYears_MP
B2F09TOB (KEEP=ID B09SMKNW B09STS3M B09SLS3M
B09CGTDY) CARDIA.C1F09TOB (KEEP=ID C09SMKNW C09STS3M C09SLS3M
C09CGTDY) CARDIA.D1F9TOBA (KEEP=ID D09SMKNW D09STS3M D09SLS3M
D09CGTDY) CARDIA.E1F09TOB (KEEP=ID E09SMKNW E09STS3M E09SLS3M
E09CGTDY) CARDIA.F1F09TOB (KEEP=ID F09SMKNW F09STS3M F09SLS3M
F09CGTDY) CARDIA.G3F09TOB (KEEP=ID G09SMKNW G09STS3M G09SLS3M
G09CGTDY) CARDIA.H3F09TOB (KEEP=ID H09SMKNW H09STS3M H09SLS3M
H09CGTDY) CARDIA.I1F09TOB (KEEP=ID I09SMKNW I09STS3M I09SLS3M
I09CGTDY) CARDIA.A4F10 (KEEP=ID A10CIGS A10TOBAC A10CIGAR A10PIPE
A10SMKLS) CARDIA.B2F10 (KEEP=ID B10CIGS B10TOBAC B10CIGAR B10PIPE
B10SMKLS) CARDIA.C1F10 (KEEP=ID C10CIGS C10TOBAC C10CIGAR C10PIPE
C10SMKLS) CARDIA.D1F10 (KEEP=ID D10CIGS D10TOBAC D10CIGAR D10PIPE
D10SMKLS) CARDIA.E1F10 (KEEP=ID E10CIGS E10TOBAC E10CIGAR E10PIPE
E10SMKLS) CARDIA.F1F10 (KEEP=ID F10CIGS F10TOBAC F10CIGAR F10PIPE
F10SMKLS) CARDIA.G3F10 (KEEP=ID G10CIGS G10TOBAC G10CIGAR G10PIPE
G10SMKLS) CARDIA.H3F10 (KEEP=ID H10CIGS H10TOBAC H10CIGAR H10PIPE
H10SMKLS) F3REF AGE
CARDIA.I1F10 (KEEP=ID I10CIGS I10TOBAC I10CIGAR I10PIPE
I10SMKLS) ;

BY ID;

* THIS PARTICIPANT WITHDREW CONSENT TO ALL CARDIA DATA ;
IF ID=203323105195 THEN DELETE;
AGESMK=EXAMAGE-A09SMKYR;
LABEL AGESMK='AGE AT START OF SMOKING';
RUN;

* FOR MISSING DURATION OF SMOKING AT THE FIRST EXAM ;
* IF FIRST NON-MISSING CIGS/DAY IS NON-ZERO COMPUTE AVERAGE AGE AT START OF SMOKING
AT BASELINE ;
SmokingPackYears_MP

PROC MEANS NOPRINT DATA=TEMP;
   VAR AGESMK;
   OUTPUT OUT=MEANOUT MEAN=MAGESMK;
RUN;

* LIFETIME PACK-YEARS OF CIGARETTES SMOKED;

DATA DV.PackYrs_MP (KEEP=ID PY0 PY2 PY5 PY7 PY10 PY15 PY20 PY25 PY30 LTSMK0 LTSMK2
               LTSMK5 LTSMK7 LTSMK10 LTSMK15 LTSMK20
               LTSMK25 LTSMK30);
   SET TEMP;
   IF _N_=1 THEN
      SET MEANOUT;
   ARRAY CIGS{8} B10CIGS C10CIGS D10CIGS E10CIGS F10CIGS G20CIGS H20CIGS
               I10CIGS;
   ARRAY SMK{8} B09SMKNW C09SMKNW D09SMKNW E09SMKNW F09SMKNW G09SMKNW
              H09SMKNW I09SMKNW;
   ARRAY STS{8} B09STS3M C09STS3M D09STS3M E09STS3M F09STS3M G09STS3M
              H09STS3M I09STS3M;
   ARRAY SLS{8} B09SLS3M C09SLS3M D09SLS3M E09SLS3M F09SLS3M G09SLS3M
              H09SLS3M I09SLS3M;
   ARRAY TDY{9} A09CGTDY B09CGTDY C09CGTDY D09CGTDY E09CGTDY F09CGTDY
             G09CGTDY H09CGTDY I09CGTDY;
   ARRAY EXAMS{9} _TEMPORARY_ (0, 2, 5, 7, 10, 15, 20, 25, 30);

   ARRAY F10{8, 4} B10TOBAC B10CIGAR B10PIPE B10SMKL
            C10TOBAC C10CIGAR C10PIPE C10SMKL
            D10TOBAC D10CIGAR D10PIPE D10SMKL
            E10TOBAC E10CIGAR E10PIPE E10SMKL
            F10TOBAC F10CIGAR F10PIPE F10SMKL
            G10TOBAC G10CIGAR G10PIPE G10SMKL
            H10TOBAC H10CIGAR H10PIPE H10SMKL
            I10TOBAC I10CIGAR I10PIPE I10SMKL
            ;

   * SET 0 FOR SKIP-PATTERN;
   DO I=1 TO 8;
      IF CIGS{I}=2 AND SMK{I}=1 AND TDY{I+1}<0 THEN
         TDY{I+1}=0;
      IF CIGS{I}=1 AND (STS{I}=1 OR SLS{I}=1) AND TDY{I+1}<0 THEN
         TDY{I+1}=0;
      IF (F10{I, 1}=1 OR F10{I, 2}=1 OR F10{I, 3}=1 OR F10{I, 4}=1) AND
         TDY{I+1}<0 THEN
         TDY{I+1}=0;
   END;
SmokingPackYears_MP

END;

* FOR THOSE ANSWER NO TO CURRENT SMOKING, CODE CGTDY=0;
DO I=1 TO 8;
   IF SMK{I}=1 THEN
      TDY{I+1}=0;
END;

* IMPUTE BEGIN;
** SET CIG/DAY=0 IF CURRENT SMOKING=NO;
DO I=1 TO 8;
   IF SMK{I}=1 AND TDY{I+1}<0 THEN
      TDY{I+1}=0;
END;

* FOR MISSING CIGS/DAY AT THE FIRST EXAM, ASSUME IT IS THE SAME AS THE
FIRST NON-MISSING VALUE;
   IF A09CGTDY<0 AND LASTVIS>=2 THEN
      DO I=2 TO LASTVIS WHILE (A09CGTDY<0);
         IF TDY{I}>=0 THEN
            A09CGTDY=TDY{I};
      END;

* FOR MISSING DURATION OF SMOKING AT THE FIRST EXAM;
** A) IF FIRST NON-MISSING CIGS/DAY IS 0, THEN ASSUME 0 (AND DOESN'T
MATTER);
** B) IF FIRST NON-MISSING CIGS/DAY IS NON-ZERO, THEN CALCULATE AVERAGE
AGE OF SMOKING ONSET IN CARDIA FOR BASELINE SMOKERS;
** C) IF AGE IS 17, 18, OR 19 , SET A09SMKYR = 1;
   IF A09SMKYR<0 AND LASTVIS>1 THEN
      DO I=2 TO LASTVIS WHILE (A09SMKYR<0);
         IF TDY{I}=0 THEN A09SMKYR=0;
         ELSE IF TDY{I}>0 THEN
            IF EXAMAGE<=19 THEN A09SMKYR=1;
            ELSE A09SMKYR=MAX(1, EXAMAGE-MAGESMK);
      END;

END;

* FOR MISSING VALUES AT "MIDDLE" EXAMS (NOT FIRST OR LAST), ASSUME CONSTANT
CHANGE IN PACK-YEARS TO GET FROM PRE-MISSING TO POST-MISSING VALUES;
** INCORPORATE A SMOOTHER IMPUTATION FOR MISSING VALUES;
J2=1;
CONT=1;
DO I=J2 TO LASTVIS WHILE (CONT=1 AND N(OF TDY{*})>0);

* LAST NON-MISSING DATA FOR MISSING DATA INTERVAL;
DO J=I TO 8 WHILE (TDY{J}>=0);
SmokingPackYears_MP

J0=J;
YR0=EXAMS{J};
CD0=TDY{J};
END;

* FIRST NON-MISSING DATA FOR MISSING DATA INTERVAL;
IF J0<LASTVIS THEN
   DO;
      DO J=J0+1 TO LASTVIS WHILE (TDY{J}<0);
      END;
      IF J<=LASTVIS THEN
         DO;
            YR2=EXAMS{J};
            CD2=TDY{J};
            J2=J;
         END;
      ELSE
         CONT=0;
   END;

* SMOOTHER IMPUTATION;
IF YR2>YR0>=0 THEN
   DO YRM=J0+1 TO J2-1;
      IF TDY{YRM}<0 THEN
         TDY{YRM}=CD0+(CD2-CD0)*(EXAMS{YRM}-YR0)/(YR2-YR0);
      END;
   END;

* FOR MISSING VALUES AT THE "END" - ASSUME NO CHANGE IN CIGS/DAY SINCE LAST MEASUREMENT;
IF LASTVIS>1 THEN
   DO;
      IF TDY{LASTVIS}<0 THEN
         DO I=(LASTVIS-1) TO 1 BY -1 WHILE (TDY{LASTVIS}<0);
         TDY{LASTVIS}=TDY{I};
      END;
   END;

* PACK-YEARS;
ARRAY PY{9} PY0 PY2 PY5 PY7 PY10 PY15 PY20 PY25 PY30;
DO I=1 TO 9;
   IF I=1 THEN
      PY{I}=A09SMKYR*A09CGTDY/20;
   ELSE
      PY{I}=MEAN(TDY{I}, TDY{I-1})*(EXAMS{I}-EXAMS{I-1})/20;
   END;
SmokingPackYears_MP

LTSMK0=SUM(0, PY0);
LTSMK2=SUM(0, PY0, PY2);
LTSMK5=SUM(0, PY0, PY2, PY5);
LTSMK7=SUM(0, PY0, PY2, PY5, PY7);
LTSMK10=SUM(0, PY0, PY2, PY5, PY7, PY10);
LTSMK15=SUM(0, PY0, PY2, PY5, PY7, PY10, PY15);
LTSMK20=SUM(0, PY0, PY2, PY5, PY7, PY10, PY15, PY20);
LTSMK25=SUM(0, PY0, PY2, PY5, PY7, PY10, PY15, PY20, PY25);
LTSMK30=SUM(0, PY0, PY2, PY5, PY7, PY10, PY15, PY20, PY25, PY30);

LABEL LTSMK0 = 'LIFETIME PACK-YEARS OF CIGARETTES SMOKED AT YEAR 0'
  LTSMK2 = 'LIFETIME PACK-YEARS OF CIGARETTES SMOKED AT YEAR 2'
  LTSMK5 = 'LIFETIME PACK-YEARS OF CIGARETTES SMOKED AT YEAR 5'
  LTSMK7 = 'LIFETIME PACK-YEARS OF CIGARETTES SMOKED AT YEAR 7'
  LTSMK10 = 'LIFETIME PACK-YEARS OF CIGARETTES SMOKED AT YEAR 10'
  LTSMK15 = 'LIFETIME PACK-YEARS OF CIGARETTES SMOKED AT YEAR 15'
  LTSMK20 = 'LIFETIME PACK-YEARS OF CIGARETTES SMOKED AT YEAR 20'
  LTSMK25 = 'LIFETIME PACK-YEARS OF CIGARETTES SMOKED AT YEAR 25'
  LTSMK30 = 'LIFETIME PACK-YEARS OF CIGARETTES SMOKED AT YEAR 30'

  PY0 = 'PACK-YEARS AT YEAR 0'
  PY2 = 'PACK-YEARS AT YEAR 2'
  PY5 = 'PACK-YEARS AT YEAR 5'
  PY7 = 'PACK-YEARS AT YEAR 7'
  PY10 = 'PACK-YEARS AT YEAR 10'
  PY15 = 'PACK-YEARS AT YEAR 15'
  PY20 = 'PACK-YEARS AT YEAR 20'
  PY25 = 'PACK-YEARS AT YEAR 25'
  PY30 = 'PACK-YEARS AT YEAR 30'

; 

KEEP ID
  PY0 PY2 PY5 PY7 PY10 PY15 PY20 PY25 PY30
  LTSMK0 LTSMK2 LTSMK5 LTSMK7 LTSMK10 LTSMK15 LTSMK20 LTSMK25 LTSMK30;

RUN;

ODS LISTING;
PROC CONTENTS DATA = DV.PACKYRS_MP VARNUM;
TITLE "DERIVED DATA: PACKYRS_MP"
RUN;
TITLE;
ODS LISTING CLOSE;