HCHS/SOL Visit 3
DERIVED VARIABLE DICTIONARY

INV Version 2.0
September 2024

Prepared by
HCHS/SOL Coordinating Center
Collaborative Studies Coordinating Center
UNC Department of Biostatistics

Franklyn Gonzalez II
Nicole Slye
Daniela Sotres-Alvarez
Jianwen Cai

HCHS/SOL Visit 3 Derived Variable Dictionary
INV Version 2.0, September 2024

Table of Contents

10.   CLINICAL CHARACTERISTICS   41

VARIABLE MODIFICATIONS   48

Updates to HCHS/SOL Visit 3 Data Release or Documentation

Version	Date	Description	Documentation
_INV1	06/28/2023	**1st Data Release (interim data; n=7,458)**	V1.0 (May 2023)
_INV2	09/19/2024	**2nd Data Release (Final V3 data; n=9,865)** **Summary of updates (details are at the end):** - 83 new variables - 1 updated/corrected variable	V2.0 (September 2024)

1. DESIGN

**For the following Sample Weight variable(s): Please see HCHS/SOL Manuscript "Sample design and cohort selection in the Hispanic Community Health Study/Study of Latinos" (Annals of Epidemiology, Volume 20, Issue 8, September 2010, Pages 629-641) for information on how the sample weights were created. Also see the HCHS/SOL Analyses Methods at Baseline (Sept. 2016) and Analyses Methods for Visit 2 (March 2018) documentation that describes the proper use of sample weights in statistical analyses.

As in any complex survey design, and as done in HCHS/SOL baseline and Visit 2, sampling weights account for non-response. One important and big difference between non-response at baseline and subsequent visits is that at baseline all we knew from non-responders was age and sex (from screening) whereas at Visit 2 and Visit 3 we know all their baseline data. The calculation of the final sampling weights (FROZEN AS OF 9/2024) for Visit 3 is based on the nonresponse adjusted sampling weights for Visit 1 and taking into account the participant nonresponse for Visit 3. See the Analyses Methods for Visit 3 (September 2024) document for more information. There were eight variables used to account for non-response: Center, gender, age groups at Visit 1 (AGEGROUP_C6), AFU Refusal (Y/N), distance from clinic from recent Visit 3 address (categorized), current cigarette smoking status (CURRENT_SMOKER) and grouped categories using Cystatin C (mg/L) (LABA101), and then design stratification variable (STRAT) are used to form finer strata and the nonresponse rate for Visit 3 is then calculated for each stratum. The sampling weights are calculated based on Visit 1 nonresponse sampling weights and these nonresponse rates for Visit 3. The sampling weights for V3 are then trimmed (to handle extreme values), calibrated (to the US 2010 Census population in the target population census tracks), and normalized (so that the sum of the sampling weights adds to the total sample size).

1.1 WEIGHT_ EXPANDED_ALL_V3 (HCHS V3 Expanded Census 2010 Calibrated, Trimmed, Nonresponse Adjusted Weights (frozen as of 9/2024))

The expanded sample weight is calculated from the Visit 1 sampling weights and nonresponse rates for Visit 3, and then trimmed and calibrated.

1.2 WEIGHT_ NORM_OVERALL_ALL_V3 (HCHS V3 Overall Normalized, Census 2010 Calibrated, Trimmed, Nonresponse Adjusted Final Weights, In-Person Exam or Phone Only (frozen as of 9/2024))

The normalized sample weight is normalized to the overall sample size including participants who only had a Visit 3 phone interview and not an in-person exam. **This final sample weight variable is used for MOST analyses.**

1.3 WEIGHT_NORM_CENTER_ALL_V3 (HCHS V3 Center Normalized, Census 2010 Calibrated, Trimmed, NonResponse Adjusted Weights (frozen as of 9/2024))

ONLY use this sampling weight for analyses for which the population of interest is only ONE field center.

1.4 WEIGHT_ EXPANDED_EXAMONLY_V3 (HCHS V3 Expanded Census 2010 Calibrated, Trimmed, Nonresponse Adjusted Weights (frozen as of 9/2024))

The expanded sample weight is calculated from the Visit 1 sampling weights and nonresponse rates for Visit 3, and then trimmed and calibrated.

1.5 WEIGHT_ NORM_OVERALL_EXAMONLY_V3 (HCHS V3 Overall Normalized, Census 2010 Calibrated, Trimmed, Nonresponse Adjusted Final Weights, In-Person Exam (frozen as of 9/2024))

The normalized sample weight is normalized to the overall sample size of those who attended the Visit 3 in-person exam.

1.6 WEIGHT_NORM_CENTER_EXAMONLY_V3 (HCHS V3 Center Normalized, Census 2010 Calibrated, Trimmed, NonResponse Adjusted Weights (frozen as of 9/2024))

ONLY use this sampling weight for analyses for which the population of interest is only ONE field center.

2. ADMINISTRATIVE

2.1 CONSENT_V3 (Informed consent at Visit 3)

This is a binary variable that determines whether a study participant consented to participate in the Visit 3 examination study.

If Visit 3 ICT0d in (1,2) then CONSENT_V3=1;
Else if ICT0d=0 then CONSENT_V3=0;
Else CONSENT_V3=. ;

Source variable(s):
ICT0d. Agrees to participant in HCHS/SOL study (yes/no)

2.2 PARTICIPANT_EXAMONLY_V3 (Participants with in-person exam component)

This is a binary variable that determines whether or not a study participant completed the Visit 3 in-person examination component. Participants with partial interviews must have at least one of the following procedures to be considered an in-person examination: Sitting Blood Pressure, Anthropometry, or Lab.

PARTICIPANT_EXAMONLY_V3= (CONSENT_V3=1 & CHK22 in (1,5) & ELE3=5) OR (CONSENT_V3=1 & CHK22 =2 & ELE3=5 & cmiss(SBP1, ANT1, LABA66, LABA68, LABA69, LABA70, LABA72, LABA107)< 8)

Source variable(s):
CONSENT_V3. Informed consent at V3
CHK22. V3 Exam summary
ELE3. Individual Participation Status
SBP1. Arm used for sitting blood pressure
ANT1. Able to stand
LABA66. Total Cholesterol (mg/dL)
LABA68. HDL - Cholesterol (mg/dL)
LABA69. LDL - Cholesterol (mg/dL)
LABA70. Fasting glucose (mg/dL)
LABA72. % Glycosylated Hemoglobin (A1C)
LABA107. Triglycerides (mg/dL)

2.3   PHONE_EXAM_CONVERT_V3 (Converted from phone only to in-person exam)

This is a binary variable that determines if a participant started Visit 3 as a phone interview only and then later completed the in-person exam component. This variable is missing for participants who did not start Visit 3 as a phone interview and only completed the in-person exam component.

If Visit 3 ICT0c = '3' and ICT0d in (1,2) for one occurrence and if ICT0c in ('1', '2') and ICT0d in (1,2) for a later occurrence then PHONE_EXAM_CONVERT_V3 = 1;

Source variable(s):
ICT0c. Consent, Update, or Phone Interview
ICT0d. Agrees to participant in HCHS/SOL study (yes/no)

2.4   VISIT_TYPE_V3 (Visit 3 visit types: 1-Clinic, 2-Home, 3-Phone only)

This documents the type of Visit 3 visit: clinic, home, or phone interview only. Participants with partial interviews must have at least one of the following procedures to be considered an in-person examination: Sitting Blood Pressure, Anthropometry, or Lab.

If CHK22 = 1 OR (CHK22 = 2 AND cmiss(SBP1, ANT1, LABA66, LABA68, LABA69, LABA70, LABA72, LABA107)< 8) then VISIT_TYPE_V3=1;
Else if CHK22=5 then VISIT_TYPE_V3=2;
Else if CHK22=6 OR (CHK22=2 & cmiss(SBP1, ANT1, LABA66, LABA68, LABA69, LABA70, LABA72, LABA107) = 8) then VISIT_TYPE_V3=3;

Response Format: 1=Clinic
                     2=Home
                     3=Phone only

Source variable(s):
CHK22. V3 Exam checklist
SBP1. Arm used for sitting blood pressure
ANT1. Able to stand
LABA66. Total Cholesterol (mg/dL)

LABA68. HDL - Cholesterol (mg/dL)
LABA69. LDL - Cholesterol (mg/dL)
LABA70. Fasting glucose (mg/dL)
LABA72. % Glycosylated Hemoglobin (A1C)
LABA107. Triglycerides (mg/dL)

2.5 HOMEVISIT_V3 (Visit 3 home visit)

This is a binary variable that determines if a participant's Visit 3 examination was completed at home.

If CHK22=5 then HOMEVISIT _V3=1;
Else HOMEVISIT_V3=0;

Source variable(s):
CHK22. V3 Exam summary

2.6 CLINDATE_V3 (Visit 3 Clinic Date)

This documents the date of the participant's clinic visit. It is derived from the most common value of following dates: Clinic Exam Checklist (CHK), Anthropometry (ANT), or Sitting Blood Pressure (SBP) forms, which are procedures usually done on the day of the exam. For participants who only completed the Visit 3 phone interview, the date is determined from the most common completion date of the forms that could be administered over the phone (Table 8 in HCHS/SOL Manual of Procedures #2): Demographics (DEM), Personal Identifiers (IDE), Participant Disability Screen (PDE), Health Use (HUE), Socioeconomic Status/Occupation (SEE), and Tobacco Use (TBE).

Source variable(s):
CHK0A. Date of completion of the Clinic Exam Checklist, where CHK22 in (1,2,5)
ANT0A. Date of completion of the Anthropometry Form
SBP0A. Date of completion of the Sitting Blood Pressure Form
DEM0A. Date of completion of the Demographics Form
IDE0A. Date of completion of the Personal Identifiers Form
PDE0A. Date of completion of the Participant Disability Screen Form
HUE0A. Date of completion of the Health Use Form
SEE0A. Date of completion of the Socioeconomic Status/Occupation Form
TBE0A. Date of completion of the Tobacco Use Form

2.7 ANY_PERMIT_V3 (Any permission for public use of data, based on ICT3 and ICT5)

This is a binary variable that combines the 2 permission variables to determine whether or not a study participant gave any permission for public use of data.

IF ICT3=1 and ICT5=1 then ANY_PERMIT_V3=1

otherwise 0

Response Format: 0 = use not permitted
                          1 = permitted

Source variable(s):
ICT3. Agrees to allow HCHS/SOL and investigators HCHS/SOL works with to study samples (blood, cells and urine) in current and future research
ICT5. Agrees to allow HCHS/SOL and investigators HCHS/SOL works with to use genetic material (DNA/RNA) in current and future research

2.8   DURATION_V3 (Difference between Arrival Time and Exit Exam Time (hrs))

This is a continuous variable that calculates the total exam duration (in hrs) using the start and end time from the Clinic Exam Checklist form (CHK) only for participants with procedures conducted at the clinic or a home visit.

DURATION_V3 =(CHK21-CHK2)/60/60;

Source variable(s):
CHK2. Arrival time
CHK21. Exit time

2.9   YRS_BTWN_V1V3 (Elapsed time in years between visits 1 and 3)

This is a continuous variable with total elapsed time in years between HCHS/SOL Visit 1 and HCHS/SOL Visit 3 clinic dates.

YRS_BTWN_V1V3 = (CLINDATE_V3-CLINDATE)/365.25

Source variable(s):
CLINDATE. Visit 1 Clinic Date
CLINDATE_V3. Visit 3 Clinic Date

2.10   YRS_BTWN_V2V3 (Elapsed time in years between visits 2 and 3)

This is a continuous variable with total elapsed time in years between HCHS Visit 2 and HCHS Visit 3 clinic dates.

YRS_BTWN_V2V3 = (CLINDATE_V3-CLINDATE_V2)/365.25

Source variable(s):
CLINDATE_V2. Visit 2 Clinic Date
CLINDATE_V3. Visit 3 Clinic Date

3. SOCIO-DEMOGRAPHIC

3.1 AGE_V3 (Age at Visit 3 in years)

This is the age of the participant in years (an integer variable) at the time of the participant's V3 clinic visit. It is determined from the participant's date of birth (confidential variable not released) and the V3 clinic visit date. If a participant's age was greater than 89 years then their age was censored to 89 (n=2).

AGE_V3 = INTEGER of (CLINDATE_V3– DOB_V3)/365.25

Source variable(s):
CLINDATE_V3. Visit 3 Clinic Date
DOB_V3. Date of birth (DOB) based on V3 forms DEM2 and PIEA2. DEM2 was prefilled with DOB from last visit available and corrected if errors were identified. NOTE that because DOB errors were identified at Visits 2 and 3, it is possible to have inconsistent ages across visits. Analysts need to correct previous ages if all three are used in the analyses.

3.2 AGEGRP_C7_V3 (7-level grouped age of participant at Visit 3)

This is the categorical (grouped) age of the participant. It is determined from the derived variable AGE_V3.

Response Format: 1=Ages 20-29
2=Ages 30-39
3=Ages 40-49
4=Ages 50-59
5=Ages 60-69
6=Ages 70-79
7=Ages 80+

Source variable(s):
AGE_V3. Age at Visit 3 in years

3.3 Combined Recoded SGE Form and SME/SOE Form Variables

Except for SEX_V3, it is strongly recommended to consult with Drs. Krista Perreira and/or Tonia Poteat on the correct use of terms and the appropriate analyses for variables about sexuality and gender identity (sections 3.3 to 3.16). They have the expertise and were the original investigators funded to collect and derive these variables as part of the SGM ancillary study (R01HL149778).

The 5-item V3 NIMHD Sexual and Gender Identity (SGE) form was replaced with the SME and SOE forms on 02/18/2021, and only Bronx had started (N=59). The old SGE form is only used when SME data is not available and will NOT be distributed. To create common variables across the two forms, we programed new variables using the new form name

and the suffix "R" to indicate recode. The following table provides a concordance between variable names in the SGE form (version 1, 3/3/2020) and the SME/SOE forms.

Concept	SGE Form	SME/SOE Form	Recoded Variable
Sex Assigned at Birth	SGE103	SME103	SME103R
Gender Identity	--	SME104	SME104
Intersex/DSD	--	SME106	SME106
Same Sex Attraction	--	SME107	SME107
Sexual Orientation Identity	SGE104	SOE110	SOE110R
Transgender Identity	SGE105	SOE111	SOE111R
Sexual Behavior Past 12 months	SGE106	SOE112	SOE112R
Sexual Attraction	SGE107	SOE108	SOE108R
Romantic Partners	--	SOE109	SOE109

Note: SGE105 (age first identify with current gender) is a follow-up question to SGE104.

3.3.1   SME103R (Sex assigned at birth recoded)

This variable combines SGE103 and SME103, recodes female as zero, and recodes don't know and refused as missing.

If SME103=1 or SGE103=1 then SME103R = 1;
If SME103=2 or SGE103=2 then SME103R = 2;

Response Format: 0=Female
                   1=Male

Source variable(s):
SME103, SGE103. Sex assigned at birth collected at V3

3.3.2   SOE108R (Sexual attraction recoded)

This variable combines SGE107 and SOE108 and recodes don't know and refused as missing.

Detailed Description:
If SGE107=1 or SOE108=1 then SOE108R=1;
If SGE107=2 or SOE108=2 then SOE108R=2;
If SGE107=3 or SOE108=3 then SOE108R=3;
If SGE107=4 or SOE108=4 then SOE108R=4;
If SGE107=5 or SOE108=5 then SOE108R=5;
If SGE107=6 or SOE108=6 then SOE108R=6;
Else SOE108R = .;

Response Format: 1=Only attracted to females, never males
                   2=Mostly attracted to females, and at least once to a male

3=Attracted about equally often to females and to males
4=Mostly attracted to males and at least once to a female
5=Only attracted to males, never females
6=Never felt sexually attracted to anyone at all
.=missing/don't know/not sure/refused

Source variable(s):
SGE107, SOE108. Sexual attraction

3.3.3 SOE110R (Sexual orientation identity recoded)

This variable combines SGE104 and SOE110 and recodes don't know and refused as missing.

Detailed Description:
If SGE104=1 or SOE110=1 then SOE110R=1;
If SGE104=2 or SOE110=2 then SOE110R=2;
If SGE104=3 or SOE110=3 then SOE110R=3;
If SGE104=4 or SOE110=4 then SOE110R=4;
If SGE104=5 or SOE110=5 then SOE110R=5;
Else SOE110R=.;

Response format: 1=Straight/Heterosexual
                    2=Lesbian
                    3=Gay
                    4=Bisexual
                    5=Other
                    .=missing/don't know/not sure/refused
Source variable(s):
SGE104, SOE110.

3.3.4 SOE111R (Transgender identity recoded)

This variable combines SGE105 and SOE111 and recodes don't know and refused as missing. This variable is not released due to small cell sizes.

If SGE105=1 or SOE111=1 then SOE111R=1;
If SGE105=2 or SOE111=2 then SOE111R=2;
If SGE105=3 or SOE111=3 then SOE111R=3;
If SGE105=4 or SOE111=4 then SOE111R=4;

Response format: 1=No
                    2=Yes, transgender, male-to-female
                    3=Yes, transgender, female-to-male
                    4=Yes, transgender, gender nonbinary
                    .=missing/dk/ref

Source variable(s):

SGE105, SOE111. Transgender identity

3.3.5   SOE112R (Gender of whom the participant had sex with in the past 12 months recoded)

This variable combines SGE106 and SOE112. This variable is not released due to small cell sizes; cell counts are displayed in the codebook.

If SGE106=1 or SOE112=1 then SOE112R=1;
If SGE106=2 or SOE112=2 then SOE112R=2;
If SGE106=3 or SOE112=3 then SOE112R=3;
If SGE106=4 or SOE112=4 then SOE112R=4;
If SOE112=5 then SOE112R=5;

Response format: 1=A man or men only
                    2=A woman or women only
                    3=Both man or men and woman or women
                    4=I have not had sex
                    5=Other genders

Source variable(s):
SGE106, SOE112. Gender of whom the participant had sex with in the past 12 months

3.4   SEX_V3 (Sex Assigned at birth updated at Visit 3)

This is a dummy variable with participant's sex assigned at birth that corrects, when necessary, baseline information using information from Visit 2 or Visit 3. At Visit 1, the PIE form (Personal Information) collected PIE1 "GENDER" but it did not distinguish between sex at birth and gender identity as it was not asked to the participant. Instead, as noted in the V1 MOP, gender may have been assigned by interviewers based on the participant's appearance. At Visit 3, sex assigned at birth was asked directly to the participant and collected first on the 5-item V3 NIMHD Sexual and Gender Identity (SGE) form which was replaced with the SME form on 02/18/2021. If SGE and SME data are not available, then information from Visit 2 or Visit 1 is used.

  If SME103R = 0 then SEX_V3 = 0;
  Else if SME103R = 1 then SEX_V3 =1;
  Else if DEM1 = 2 then SEX_V3 = 0;
  Else if DEM1 = 1 then SEX_V3 = 1;

Response Format: 0=Female
                   1=Male

Source variable(s):
DEM1. Sex prefilled in CDART demographics form using sex updated in Visit 2 when available, or from baseline.
SME103R. Sex assigned at birth collected at V3 recoded

GENDERNUM_V2. Sex updated in Visit 2 if needed
GENDERNUM. Sex (PIE1) from HCHS/SOL baseline

3.5   SEX_C3_V3 (Sex assigned at birth including intersex)

Currently, medical providers in the U.S are not provided the option of identifying a newborn as intersex or differences in sex development (DSD) on their birth certificate. However, many U.S states have allowed individuals to change their sex on a birth certificate and countries worldwide are increasingly recognizing nonbinary sexes. This variable integrates intersex or DSD and sex assigned at birth into a single variable.

IF SEX_V3 = 0 and (SME106=0 or SME106=8 or SME106 is missing) then SEX_C3 = 0;
ELSE IF SEX_V3 = 1 and (SME106=0 or SME106=8 or SME106 is missing) then SEX_C3 = 1;
ELSE IF SME106=1 then SEX_C3 = 2;

Response Format: 0=Female
                   1=Male
                   2=Intersex

Source variable(s):
SEX_V3. Sex Assigned at birth updated at Visit 3
SME106. DSD/intersex condition

3.6   INTERSEX (Intersex condition declared by a doctor)

This variable documents if an individual has been told by a doctor that they have a difference of sexual development or intersex condition.

If SME106 = 0 or SME106 = 8 or missing(SME106) then INTERSEX = 0;
Else if SME106 = 1 then INTERSEX = 1;

Source variable(s):
SME106. DSD/intersex condition

3.7   Sexual Orientation and Gender Identity Variables

Sexual orientation and gender identity variables are defined for all Visit 3 participants. Every participant has variables defined for sex assigned at birth, gender identity, sexual attraction, sexual orientation identity, and sexual behavior. It is important to remember that these variables can interact and overlap in many ways. For example, an individual whose sex assigned at birth differs from their gender identity may not identify as transgender. And an individual who identifies as transgender may have any sexual orientation identity (e.g., heterosexual, gay, lesbian, or bisexual).

3.7.1 SGM (Sexual or gender minority)

This is a variable identifying Visit 3 participants who are considered a sexual or gender minority (SGM) or not. Individuals are identified as Sexual and Gender minorities from the SGE form or the combined SME/SOE form which replaced the SGE form on 02/18/2021.

The old SGE form is utilized to determine SGM status *only* in cases where the revised SME/SOE form is not available. The 5-item NIMHD SGE form used in initial interviews does not allow for a comprehensive determination of SGM status. It erroneously excludes individuals who are gender non-binary, who have a difference of sexual development (DSD) or are intersex, and who do not identify as transgender.

Individuals identified as Sexual and Gender Minorities include those who are males (defined by sex assigned at birth) attracted to males; females(defined by sex assigned at birth) attracted to females; individuals attracted to both men and women; transgender male; transgender female; transgender non-binary; identify as lesbian, gay, bisexual, or other; males who have had sex with males in the past 12 months; and females who have had sex with females in the past 12 months; individuals who have had sex in the past 12-months with both men and women. Individuals are also identified as Sexual and Gender Minorities if their sex assigned at birth differs from their gender identity, their gender identity is non-binary or some other identity; they have been told by a doctor that they have a difference of sexual development (DSD) or an intersex condition; or they have ever felt sexually attracted to someone of the same sex.

The participant can be considered a member of a sexual or gender minority population if any of the following is true:
- SEX_V3 ≠ SME104
- SME104=3, 4
- SME106=1
- SME107=1
- SOE108R=2,3,4
- SEX_V3 =1 and SOE108R=5
- SEX_V3 =0 and SOE108R=1
- SOE109=3
- SEX_V3 =1 and SOE109=1
- SEX_V3 =0 and SOE109=2
- SOE110R=2,3,4,5
- SOE111=2,3,4
- SOE112R=3
- SEX_V3 =1 and SOE112R=1
- SEX_V3 =0 and SOE112R=2

Response Format: 0 = Non-SGM
                      1 = SGM
                      . = Missing / Don't know

Source variable(s):
SEX_V3. Sex assigned at birth
SME104. Current gender

SME106. DSD/intersex condition
SME107. Attracted to same sex
SOE108R. Sexual attraction recoded
SOE109. Romantic partners
SOE110R. Sexual orientation recoded
SOE111R. Consider themselves transgender recoded
SOE112R. Gender of whom the participant had sex with in the past 12 months recoded

Reference(s):
Center of Excellence for Transgender Health (Herman 2014)
United States Transgender Survey (James 2017)
National Survey of Family Growth (Badgett 2009, Ridolfo 2012)
National Health and Social Life Survey (Patterson 2017)

3.7.2 TRANSGENDER (Transgender)

This variable documents if an individual identifies as transgender, is male assigned at birth with a female gender identity, or is female assigned at birth with a male gender identity.

If SOE111R in (2,3,4) or (SEX_V3=1 and SME104=2) or (SEX_V3=0 and SME104=1) then TRANSGENDER = 1;
Else TRANSGENDER = 0;
If SOE111R is missing and at least one of (SEX_V3 and SME104) is missing then TRANSGENDER = .;

Source variable(s):
SOE111R. Consider self transgender recoded
SEX_V3. Sex Assigned at birth updated at Visit 3
SME104. Current gender

3.7.3 CISGENDER (Cisgender)

This is a two-category variable to indicate cisgender (1) or not (0). Cisgender refers to a person whose gender identity aligns with their assigned sex at birth. Individuals who are not cisgender may be transgender or gender non-binary. Individuals who are intersex may be cisgender or not. Sexual orientation identity can vary across all genders and among persons who are intersex or have DSD conditions. Thus, this variable may be used in combination with the 5-category variable on sexual attraction/orientation identity. There are 51 participants missing this variable because they completed the SGE form which did not ask about current gender identity.

If SOE111R in (2,3,4) or (SEX_V3=1 and SME104=2) or (SEX_V3=0 and SME104=1) then TRANSGENDER = 1;
Else TRANSGENDER = 0;
If SOE111 is missing and at least one of (SEX_V3 and SME104) is missing then TRANSGENDER = .;

Source variable(s):

SOE111R. Consider self transgender recoded
SEX_V3. Sex Assigned at birth updated at Visit 3
SME104. Current gender
TRANSGENDER. Transgender

3.7.4 GENIDENTITY_C4 (Self-reported gender identity (4 categories))

This categorical variable determines self-reported gender identities. Transgender women whose gender identity is women are categorized as women. Individuals are only categorized as transgender or nonbinary when this is their chosen gender identity or their gender identity is not available and they have indicated a transgender or non-binary identity.

If SME104=3 or SOE111R in (2,3,4) then GENIDENTITY_C4 = 3;
Else if SME104=1 then GENIDENTITY_C4 = 1;
Else if SME104=2 then GENIDENTITY_C4 = 2;
Else if SOE111R = 1 and SME104 is missing then GENIDENTITY_C4 = .;
Else GENIDENTITY_C4 = 4;

Response format: 1=Man
                    2=Woman
                    3=Gender Non-Binary or Transgender
                    4=Other

Source variable(s):
SME104. Current gender
SOE111R. Consider self transgender

3.7.5 GENIDENTITY_C7 (Self-reported gender identity (7 categories))

This categorical variable determines self-reported gender identities based on sex assigned at birth, gender identity, and transgender identity.

If SEX_V3=1 and SME104=1 then GENIDENTITY_C7 = 1;
Else if (SEX_V3=0 and SME104=1) OR (TRANSGENDER=1 and SME104=1) then GENIDENTITY_C7 = 2;
Else if SME104=3 or SOE111R=4 then GENIDENTITY_C7 = 3;
Else if (SEX_V3=1 and SME104=2) OR (TRANSGENDER=1 and SME104=2) then GENIDENTITY_C7 = 4;
Else if SEX_V3=0 and SME104=2 then GENIDENTITY_C7 = 5;
Else if SME104=4 then GENIDENTITY_C7 = 6;
Else GENIDENTITY_C7 = 9;

Response format: 1=Cis Man
                    2=Trans Man
                    3=Non-Binary
                    4=Trans Woman

5=Cis Woman
6=Other
9=Missing/Don't Know

Source variable(s):
SEX_V3. Sex assigned at birth updated at Visit 3
SME104. Current gender
TRANSGENDER. Transgender
SOE111R. Consider self transgender recoded

3.7.6   TNB (Transgender or nonbinary)

This is an indicator variable to identify an individual who identifies as transgender or nonbinary.

IF SME104=3 or SOE111R in (2,3,4) then TNB = 1;
ELSE TNB = 0;
If all source variables are missing then TNB = .;

Source variable(s):
SME104. Current gender
SOE111R. Consider self transgender recoded

3.7.8   TGD (Trans and gender diverse)

This variable identifies someone who is trans and gender diverse (TGD). This means they are either transgender, nonbinary, intersex, some other identity, or some combination based on their identities or on the difference between assigned sex at birth and gender identity.

If TRANSGENDER = 1 OR INTERSEX = 1 OR TNB = 1 OR SME104 = 4 then TGD=1;
Else TGD=0;
If all source variables are missing then TGD = .;

Source variable(s):
TRANSGENDER. Transgender
INTERSEX. Intersex condition declared by a doctor
TNB. Transgender or nonbinary
SME104. Current gender

3.7.9   LGBTID (Lesbian, gay, bisexual, nonbinary, or transgender)

This variable indicates whether a participant identifies as lesbian, gay, bisexual, nonbinary, or transgender. The variable is based on identity only and not on sexual attraction or behavior.

If SOE110R in (2,3,4) or (SOE110R in (5,8) and SOE110a = 1) or SOE111R in (2,3,4) or SME104 = 3 then LGBTID = 1;
Else LGBTID = 0;
If all source variables are missing then LGBTID = .;

Source variable(s):
SOE110R. Sexual orientation recoded
SOE110a. Sexual orientation supplement
SOE111R. Consider self transgender recoded
SME104. Current gender

3.7.10 SXORIENTATION (Sexual orientation identity or sex assigned at birth and self-reported sexual attraction)

This categorical variable identifies sexual orientation based on sexual orientation identity or sex assigned at birth and self-reported sexual attraction. This variable does not use sexual behavior. Sex assigned at birth is used in this classification because gender identity was not available for individuals completing SGE forms. In this variable, an individual who identifies as heterosexual but who is assigned male at birth and attracted only to men would be classified as male attracted to men; an individual who identifies as heterosexual but who is assigned female at birth and attracted only to females would be classified as female attracted to women.

Participants are only attracted to the "opposite sex" or straight/heterosexual if:
    SOE110R=1 OR
    SEX_V3=1 and SOE108R=1 OR
    SEX_V3=0 and SOE108R=5 OR
    SME107 = 0 and SOE108R is missing and SOE110R is missing
Participants are men attracted to men or gay if:
    SOE110R=3 OR
    SEX_V3=1 and SOE108R=5 OR
    SEX_V3=1 and SME107=1
Participants are women attracted to women or lesbian if:
    SOE110R=2 OR
    SEX_V3=0 and SOE108R=1 OR
    SEX_V3=0 and SME107=1
Participants are attracted to both men and women or bisexual if:
    SOE108R in (2,3,4) OR
    SOE110R=4
Participants not classified above are classified as other/missing/don't know

Response format: 0=Only Attracted to the opposite sex or Straight/ Heterosexual
                   1=Men Attracted to Men or Gay
                   2=Women Attracted to Women or Lesbian
                   3=Attracted to both Men and Women or Bisexual
                   9=Other/Missing/Don't Know

Source variable(s):
SOE110R. Sexual orientation recoded
SEX_V3. Sex Assigned at birth updated at Visit 3
SOE108R. Sexual attraction recoded
SME107. Attracted to same sex

3.7.11 SAMESXATT (Ever been attracted to the same sex)

This variable indicates whether a participant has ever been attracted to the same sex.

If SOE108R in (2,3,4) or (SME107=1) or (SEX_V3=1 and SOE108R=5) or (SEX_V3=0 and SOE108R=1) then SAMESXATT = 1;
Else SAMESXATT = 0;
If SME107 is missing and at least one of (SEX_V3 and SOE108R) is missing then SAMESXATT = .;

Source variable(s):
SOE108R. Sexual attraction recoded
SME107. Attracted to same sex
SEX_V3. Sex Assigned at birth updated at Visit 3

3.7.12 SAMESXPART (Has had a same sex romantic or sexual partner)

This variable identifies whether a participant has had a same sex romantic or sexual partner. Individuals who indicate they have romantic partner whose gender is "other" are not included as having a same sex partner.

If SOE109=3 or (SEX_V3=1 and SOE109=1) or (SEX_V3=0 and SOE109=2) or SOE112R=3 or (SEX_V3=1 and SOE112R=1) or (SEX_V3=0 and SOE112R=2) then SAMESXPART = 1;
Else SAMESXPART = 0;
If all source variables are missing then SAMESXPART = .;

Source variable(s):
SOE109. Romantic partners
SEX_V3. Sex Assigned at birth updated at Visit 3
SOE112R. Gender of whom the participant had sex with in the past 12 months recoded

3.7.13 OPPSXPART (Has had an opposite sex romantic or sexual partner)

This variable identifies whether a participant has had an opposite sex romantic or sexual partner. Individuals who indicate they have romantic partner whose gender is "other" are not included as having an opposite sex partner.

If SOE109=3 or (SEX_V3=1 and SOE109=2) or (SEX_V3=0 and SOE109=1) or SOE112R=3 or (SEX_V3=1 and SOE112R=2) or (SEX_V3=0 and SOE112R=1) then OPPSXPART = 1;
Else OPPSXPART= 0;
If all source variables are missing then OPPSXPART = .;

Source variable(s):
SOE109. Romantic partners
SEX_V3. Sex Assigned at birth updated at Visit 3
SOE112R. Gender of whom the participant had sex with in the past 12 months recoded

3.7.14 ASEXUAL (Has never felt sexually attracted to anyone)

This variable identifies someone who has never felt sexually attracted to anyone.

If SOE108R=6 then ASEXUAL = 1;
Else if SOE108R in (1,2,3,4,5,8) then ASEXUAL = 0;
If all source variables are missing then ASEXUAL = .;

Source variable(s):
SOE108R. Sexual attraction recoded

3.7.15 AGE_GENDERID_LE_AGE_V1 (Age for gender identity less than or equal to age at visit 1)

This binary variable indicates if the age at which someone first identified with their current gender is less than or equal to their age at HCHS Visit 1.

If SME105 is not missing AND SME105 <= AGE then AGE_GENDERID_LE_AGE_V1 = 1;
Else if AGE is not missing AND SME105 > AGE then AGE_GENDERID_LE_AGE_V1 = 0;

Source variable(s):
SME105. Age identify current gender
AGE. Age at Visit 1

3.7.16 AGE_SOID_LE_AGE_V1 (Age for sexual orientation identity less than or equal to age at visit 1)

This binary variable indicates if the age at which someone first felt attracted to someone of the same sex less than or equal to their age at HCHS Visit 1. This variable is missing for participants who have never felt sexually attracted to someone of the same sex.

If SME107A is not missing AND SME107A <= AGE then AGE_SOID_LE_AGE_V1 = 1;
Else if AGE is not missing AND SME107A > AGE then AGE_SOID_LE_AGE_V1 = 0;

Source variable(s):
SME107A. Age attracted to same sex
AGE. Age at Visit 1

3.8 INCOME2_V3 (Grouped yearly household income at Visit 3)

This derived variable groups each subject to one income category. The income is set to missing if SEE1 and SEE1a are both missing. Note these categories differ from INCOME_V2 at Visit 2 due to differences in the SEE income questions.

If .z <SEE1<999999 then do;
   If SEE1<10000 then INCOME2_V3=1;
   Else if SEE1<=15000 then INCOME2_V3=2;
   Else if SEE1<=20000 then INCOME2_V3=3;
   Else if SEE1<=25000 then INCOME2_V3=4;
   Else if SEE1<30000 then INCOME2_V3=5;
   Else if SEE1<=40000 then INCOME2_V3=6;
   Else if SEE1<=50000 then INCOME2_V3=7;
   Else if SEE1<=70000 then INCOME2_V3=8;
   Else if SEE1<=100000 then INCOME2_V3=9;
   Else INCOME2_V3=10; End;
Else if .z <SEE1a<99 then do;
   If SEE1a in (1 2 3 4 5) then INCOME2_V3=SEE1a;
   Else if SEE1a in (6 7) then INCOME2_V3=6;
   Else if SEE1a in (8 9) then INCOME2_V3=7;
   Else if SEE1a in (10 11 12 13) then INCOME2_V3=8;
   Else if SEE1a in (14 15) then INCOME2_V3=9;
   Else INCOME2_V3=10; End;
Else INCOME2_V3=.

Response format: 1=Less than $10,000
                    2=$10,001-$15,000
                    3=$15,001-$20,000
                    4=$20,001-$25,000
                    5=$25,001-$29,999
                    6=$30,000-$40,000
                    7=$40,001-$50,000
                    8=$50,001-$70,000
                    9=$70,001-$100,000
                    10=More than $100,000

Source variables:
SEE1. What was your family (household) income?
SEE1a. About how much was your current family (household) income?

3.9   INCOME2_C3_V3 (3-level grouped yearly household income at Visit 3)

This derived variable groups each subject to one income category based on SEE1 and SEE1a. SEE1a was answered by some participants who refused to answer SEE1. Participants who reported their income was $0 are included in the less than $30,000 category. Participants with missing SEE1 and SEE1a are grouped together. Note these categories differ from INCOME_C3_V2 at Visit 2 due to differences in the SEE income questions.

If 0<=SEE1<999999 then do;
   If SEE1<30000 then INCOME_C3_V3=1;
   Else INCOME_C3_V3=2;
End;
Else if .z<SEE1a<6, then INCOME_C3_V3=1;
Else if 6<=SEE1a<99, then INCOME_C3_V3=2;
Else INCOME_C3_V3=3;

Response format: 1=$0 to $30,000
                    2=$30,001 or more
                    3=Missing

Source variables:
SEE1. What was your family (household) income?
SEE1a. About how much was your current family (household) income?

3.10   INCOME2_C5_V3 (5-level grouped yearly household income at Visit 3)

This derived variable groups each subject to one income category. The income is set to missing if both SEE1 and SEE1a are missing. Note these categories differ from INCOME_C3_V2 at Visit 2 due to differences in the SEE income questions.

If .z<SEE1<999999 then do;
   If SEE1<=10000 then INCOME2_C5_V3=1;
   Else if SEE1<=20000 then INCOME2_C5_V3=2;
   Else if SEE1<=40000 then INCOME2_C5_V3=3;
   Else if SEE1<=70000 then INCOME2_C5_V3=4;
   Else INCOME2_C5_V3=5; End;
Else if .z<SEE1a<99 then do;
   If SEE1a=1, then INCOME2_C5_V3=1;
   Else if SEE1a in (2, 3) then INCOME2_C5_V3=2;
   Else if SEE1a in (4, 5, 6, 7) then INCOME2_C5_V3=3.;
   Else if 40000<SEE1<=70000 then INCOME2_C5_V3=4;
   Else INCOME2_C5_V3=5; End;
Else INCOME2_C5_V3=.;

Response format: 1=Less than or equal to $10,000
                    2=$10,001-$20,000
                    3=$20,001-$40,000
                    4=$40,001-$70,000
                    5=More than $70,000

Source variables:
SEE1. What was your family (household) income?
SEE1a. About how much was your current family (household) income?

3.11 EDUCATION_C3_V3 (3-level group education level at Visit 3)

The educational status derived variable groups the attainment of a high school diploma or an equivalent degree, not having that high school or equivalent degree, or attainment of education beyond a high school equivalent (i.e. college or vocational). This variable is based on the Visit 1 (EDUCATION_C3) and Visit 2 (EDUCATION_C3_V2) responses with updates to educational attainment indicated on questions SEE10 and SEE11, as follows:

If SEE10=1 and EDUCATION_C3_V2>. then do;
If SEE11 in (4,5) then =3;
Else if SEE11 = 3 and EDUCATION_C3_V2 < 3 then =2;
Else if SEE11 in (1,2) and EDUCATION_C3_V2 < 2 then =1;
Else =EDUCATION_C3_V2;
End;
Else if SEE10=1 and EDUCATION_C3_V2=. then do;
If SEE11 in (4,5) then =3;
Else if SEE11 = 3 and EDUCATION_C3 < 3 then =2;
Else if SEE11 in (1,2) and EDUCATION_C3 < 2 then =1;
Else =EDUCATION_C3;
End;
Else if EDUCATION_C3_V2>. then =EDUCATION_C3_V2;
Else =EDUCATION_C3;

Response format:　　1=No high school diploma or GED
　　　　　　　　　　2=At most a high school diploma or GED
　　　　　　　　　　3=Greater than high school (or GED) education

Source variable(s):
EDUCATION_C3. Education Status (3 levels) at Visit 1
EDUCATION_C3_V2. Education Status (3 levels) updated at Visit 2
SEE10. Have you been involved in any educational or training program since the first SOL center visit
SEE11. What was the highest grade/level of education achieved

3.12 EMPLOYEDFY_V3 (Participant was employed for the full year at visit 3)

This binary variable indicates if a participant was employed for the previous 12 months when asked at Visit 3. Retirees who are currently working are included among those who are employed. The employed also include both full-time and part-time workers

If SEE4 = 1 and SEE5 = 12 then EMPLOYEDFY_V3 = 1;
Else EMPLOYEDFY_V3 = 0;
If SEE4 is missing then EMPLOYEDFY_V3 is missing

Source variable(s):
SEE4. Paid employment past 12 months
SEE5. Months worked past 12 months

3.13   MARITAL_STATUS_V3 (Marital Status at Visit 3)

Marital status at Visit 3 was collected in an administrative form called DEM which is not distributed. Instead, it is included in the PART_DERV_V3 file.

MARITAL_STATUS_V3 = DEM4

Response format: 1=Single
                    2=Married
                    3=Separated
                    4=Divorced
                    5=Widow
                    6=Living with a Partner

Source variable(s):
DEM4. Marital Status

3.14   HHSIZE_V3 (Household Size at Visit 3)

Based on the number of people supported by the households' income, this is an indicator variable of household size at Visit 3. It is top coded for a household size of 5 or more.

HHSIZE_V3 = SEE2
If SEE2 >=5 then HHSIZE_V3 = 5

Source variable(s):
SEE2. People supported on income

3.15   HC_V3 (Current Health Insurance Coverage at Visit 3)

This is a variable indicating current health insurance coverage (yes=1, no=0). An individual is considered currently insured if they indicate that they currently have insurance or they specify a response on type of insurance. Individuals who indicate that their only source of health insurance coverage is the Indian Health Services (IHS) are not included as being insured. IHS is not a health insurance provider and the IHS can provide healthcare to only eligible Alaska Native and American Indians at its federal hospitals and clinics. Previous versions of the health insurance coverage variable from Visit 1 (N_HC) and Visit 2 (N_HC_V2) include IHS.

HC_SUM = sum(of HCE14a - - HCE14f HCE14h)
NHC_SUM = sum(of HCE18a - - HCE18k)

If ((HCE13 is missing or in (8,9)) and HC_SUM > 0) or HCE13 = 1 then HC_V3 = 1;
Else if ((HCE13 is missing or in (8,9)) and NHC_SUM > 0) or HCE13 in (0,8,9) then HC_V3 = 0;

Response format: 0=No current health insurance
                    1=Currently have health insurance

Source variable(s):
HCE13. Do you have health insurance or other health care coverage?
HCE14a-HCE14h. What type of health insurance or health care coverage do you have?
HCE18a –HCE18k. Reasons you stopped being covered by health insurance

4. ACCULTURATION

4.1 LANG_PREF_V3 (Language preference at Visit 3)

This variable determines which language was preferred to be used for the Visit 3 examination. First, the Participant Eligibility form (ELE) is used which directly asks the participant which language they prefer to use. If this variable is missing, then three core forms are used: Participant Disability Screener (PDE), Personal Identifiers (IDE), and Medical History (MHE). A majority of those used in one language determines which language was preferred.

Response format: 1=Spanish
                   2=English

Source variable(s):
ELE1. Do you prefer to communicate in Spanish or English?
Form Variable of PDE
Form Variable of IDE
Form Variable of MHE

4.2 YRSUS_V3 (Number of years lived in the US (50 States/DC) at Visit 3)

This is a numeric variable indicating the number of years lived in the US (50 States). It uses the Visit 1 YRSUS as well as duration between V1 and V3 to update the years in US for Visit 3, as follows:

If US_BORN_V3=1 then set YRSUS_V3=AGE_V3;
Else YRSUS_V3 = YRSUS_V1+YRS_BTWN_V1V3.

Source variable(s):
US_BORN_V3. Born in US updated at Visit 3
YRSUS_V1. Years lived in the US at Visit 1
YRS_BTWN_V1V3. Elapsed time between visits 1 and 2 (yrs)

4.3 YRSUS2_C2_V3 (2-level grouped years lived in the US (50 States/DC) at Visit 3)

This is a 2-level grouped numeric variable that defined less than 20 years lived in the US versus 20 or more years lived in the US. It uses number of years lived in the US (50 States) at Visit 3, as follows:

If .Z < YRSUS_V3 < 20 then YRSUS2_C2_V3 = 1;
If .Z < YRSUS_V3 >= 20 then YRSUS2_C2_V3 = 2;
Else YRSUS2_C2_V3 =.;

Response format: 1=Less than 20 years
                    2=20 years or more

Source variable(s):
YRSUS_V3. Number of years lived in the US (50 States) at Visit 3

4.4   YRSUS2_C3_V3 (3-level grouped years lived in the US (50 States/DC) at Visit 3)

This is a 3-level grouped numeric variable that defined less than 20 years lived in the US versus 20 or more years lived in the US, plus US born group. It uses number of years lived in the US (50 States) at Visit 3 and US born information, as follows:

If US_BORN_V3 = 1 then YRSUS2_C3_V3 =3;
Else if YRSUS2_C2_V3 =2 then YRSUS2_C3_V3 =2;
Else if YRSUS2_C2_V3 =1 then YRSUS2_C3_V3 =1;

Response format: 1=Less than 20 years
                    2=20 years or more
                    3=US born

Source variable(s):
US_BORN_V3. US born updated at Visit 3
YRSUS_V3. Number of years lived in the US (50 States) at Visit 3

5.   SOCIOCULTURAL

5.1   DISCRIMEVER_V1V3 (Experienced or witnessed any discrimination in V1 or V3)

This binary variable identifies whether an HCHS/SOL participant ever reported experiencing or witnessing any discrimination in HCHS/SOL Visit 1 or Visit 3. No data on experiencing discrimination was collected in Visit 2.

If DISCRIMANY = 1 or SCEA22 in (2,3,4) or SCEA23 in (2,3,4) or SIE88 in (1,2,3) then DISCRIMEVER_V1V3 = 1;
Else if DISCRIMANY = 0 or SCEA22 = 1 or SCEA23 =1 or SIE88 = 0;

Source variable(s):
SCEA22-SCEA23. Experienced discrimination (Visit 1)
SIE88. Experienced discrimination (Visit 3)
DISCRIMANY. Experienced any discrimination (from SGM SOL Ancillary Study)

5.2   Stress of Immigration Scale

Adapted from the 21-item Stress of Immigration Survey, the 4-item Stress of Immigration Scale (SIS) in HCHS/SOL Visit 3 (SIE84-SIE87) identifies stress/worry about communicating in English, legal status documentation, taking time off work, and detention

or deportation of family members. Scores are only calculated if at least 3 items are not missing.

Reference: Sternberg et al. (2016)

5.2.1   IMMSTRESSSUM_V3 (Stress of Immigration Scale sum score)

IMMSTRESSSUM_V3 = sum(SIE84,SIE85,SIE86,SIE87);

Source variable(s):

5.2.2   IMMSTRESSAVG_V3 (Stress of Immigration Scale average score)

IMMSTRESSAVG_V3 = mean(SIE84,SIE85,SIE86,SIE87);

Source variable(s): SIE84,SIE85,SIE86,SIE87

5.3   Chronic Stress Scores

Derived from the Pilkonis Life Event Schedule and measured at Visit 3, this measure of chronic stress consists of 3 parts (did participants experience the stressor (0=No, 1=Yes); has the stressor been a problem for six months or more (0=No, 1=Yes); and severity of stressor (1=Not very stressful, 2=Moderately stressful, 3=Very stressful). The questions also cover 7 domains – personal health problem (SIE76); family health problems of close others (SIE77, SIE81), work stress (SIE78), financial stress (SIE79), relationship stress (SIE80), caregiving stress (SIE82), and other stress (SIE83).

Reference: Bromberger & Matthews (1996); Monroe & Roberts (1990); Shivpuri et al. (2012); Gallo et al. (2014); McCurley et al (2015)

5.3.1   CHRSTRTOT_V3 (Total number of ever chronic stressors present)

If at least 6 non-missing variables are present then:
CHRSTRTOT_V3 = sum(SIE76,SIE77,SIE78,SIE79,SIE80,SIE81,SIE82,SIE83);

Source variable(s): SIE76,SIE77,SIE78,SIE79,SIE80,SIE81,SIE82,SIE83

5.3.2   CHRSTRTOT6_V3 (Total number of chronic stressors present for at least 6 months, V3)

If CHRSTRTOT_V3 = 0 then CHRSTRTOT6_V3 = 0;
Else
CHRSTRTOT_V3=sum(SIE76a,SIE77a,SIE78a,SIE79a,SIE80a,SIE81a,SIE82a,SIE83b);

Source variable(s): SIE76a,SIE77a,SIE78a,SIE79a,SIE80a,SIE81a,SIE82a,SIE83b

5.3.3 CHRSTRSEV_V3 (Chronic stress severity score, V3)

If CHRSTRTOT_V3=0 then CHRSTRSEV_V3=0;
Else
CHRSTRSEV_V3=sum(SIE76b,SIE77b,SIE78b,SIE79b,SIE80b,SIE81b,SIE82b,SIE83c);

Source variable(s): SIE76b,SIE77b,SIE78b,SIE79b,SIE80b,SIE81b,SIE82b,SIE83c

5.3.4 M_V_TOT_V3 (Total Moderate to Very Chronic Stress Score, V3)

Recode (SIE Items 76b, 77b, 78b, 79b, 80b, 81b, 82b, 83c) into 1=0, 2=1, 3=1.

Add together the number of "chronic" stressors rated as moderately or very stressful. To score one point, the person would have to state that 1) yes the stressor occurred and 2) yes it lasted at least 6 months and 3) it was moderately or very stressful. Then add the number of these points and that is the total number of "moderately or very stressful chronic stressors".

M_V_TOT_V3 = Sum recoded items (SIE76b, 77b, 78b, 79b, 80b, 81b, 82b, 83c)
The score was calculated only if there is no missing items.

Source variable(s): SIE76b, 77b, 78b, 79b, 80b, 81b, 82b, 83c

5.4 Brief Resilience Scale

The Brief Resilience Scale (BRS) measures the ability to recover from stress using 6-items (SIE89-SIE94). Values should range from 1=Strongly disagree to 5=strongly agree. Higher values indicate more resilience. Three items (SIE90, 92, 94) are reverse coded so that 5=1, 4=2, 3=3, 2=4, 1=5.

References: Smith et al. (2008)

5.4.1 BRS_V3 (Brief Resilience Scale Sum)

BRS_V3 = sum(SIE89,SIE90r,SIE91,SIE92r,SIE93,SIE94r);

5.4.2 BRSAVG_V3 (Brief Resilience Average Scale)

BRS_V3 = mean(SIE89,SIE90r,SIE91,SIE92r,SIE93,SIE94r);

6.   WELL-BEING

6.1   PHQ8_V3 (PHQ-8 item measure of depressive symptoms)

This is an 8-item measure of depressive symptoms at Visit 3 calculated by summing the items (SIE95-SIE102). However, since we allow for some missing values in the data, for those observations with missing it is defined as 8 times the mean of the non-missing. Cut points on the PHQ-8 are identical to the PHQ-9. Scores of 5, 10, 15, and 20 represent cut points for mild, moderate, moderately severe and severe depression, respectively. The CESD-10 was used at Visits 1 and Visits 2. Thus, depression symptom scores cannot be used longitudinally.

Reference: Alpizar et al. (2018); Pagán-Torres et al. (2020)

  If nmiss(of SIE95 - SIE102)=0 then PHQ8_V3 = sum(of SIE95-SIE102);
  Else if nmiss(of SIE95 - SIE102)<=2 then PHQ8_V3 = 8 * mean(of SIE95-SIE102);

Source variable(s):
SIE95. Little interest or pleasure in doing things
SIE96. Feeling down, depressed, or hopeless
SIE97. Trouble falling or staying asleep, or sleeping too much
SIE98. Feeling tired or having little energy
SIE99. Poor appetite or overeating
SIE100. Feeling bad about yourself—or that you are a failure or have let yourself or your family down
SIE101. Trouble concentrating on things, such as reading the newspaper or watching television
SIE102. Moving or speaking so slowly that other people could have noticed? Or the opposite—being so fidgety or restless that you have been moving around a lot more than usual

6.2   PHQ8_GE10_V3 (Moderate/Severe Depression Symptoms (PHQ8 >= 10))

This is a measure of moderate to severe depression based on the cut-off point of 10 or greater using the PHQ8. The cut-off point is similar to the cut-off point of 10 used to measure moderate to severe depression using the CESD-10 which has a score of 0-30. Thus, this variable may be used as a longitudinal measure of moderate-severe depression in concert with CESD-based measures from Visit 1 and Visit 2.

If PHQ8_V3 >= 10 then PHQ8_GE10_V3 = 1;
Else PHQ8_GE10_V3 = 0;

Source variable(s):
PHQ8_V3. PHQ-8 item measure of depressive symptoms

7.   ANTHROPOMETRY

7.1   BMI_V3 (Body Mass Index (Kg/m2) at Visit 3)

This is a continuous measure of Body Mass Index (kg/m2) derived from participant's weight (kg) divided by height squared (m^2) at Visit 3. If either participant's weight or height is missing, the value of this variable is missing.

If ANT2 & ANT4 are not missing then BMI_V3 = ANT4 / (ANT2/100)^2

Source variable(s):
ANT2. Standing height in centimeters
ANT4. Weight in kilograms

7.2   BMIGRP_C4_V3 (4-level grouped Body Mass Index (WHO) at Visit 3)

This variable is created using the calculated body mass index to define categories based on current WHO classifications.

If 0<=BMI_V3<18.5 then BMIGRP_C4_V3=1;
Else if 18.5<=BMI_V3<25 then BMIGRP_C4_ V3=2;
Else if 25<=BMI_V3<30 then BMIGRP_C4_V3=3;
Else if 30<=BMI_V3 then BMIGRP_C4_v3=4;
Else if BMI_V3=.Z then BMIGRP_C4_V3=.;

Response format: 1=Underweight (BMI < 18.5)
                    2=Normal (18.5 ≤ BMI < 25)
                    3=Overweight (25 ≤ BMI < 30)
                    4=Obese (BMI ≥ 30)
Source variable(s):
BMI_V3. Body Mass Index (Kg/m2) at Visit 3

7.3   BMIGRP_C6_V3 (6-level grouped Body Mass Index (WHO) at Visit 3)

This variable is created using the calculated body mass index to define categories based on current WHO classifications.

If BMI_V3<=.Z then BMIGRP_C6_V3=.;
Else if 0<= BMI_V3< 18.5 then BMIGRP_C6_V3=1;
Else if 18.5<= BMI_V3< 25 then BMIGRP_C6_V3=2;
Else if 25<= BMI_V3< 30 then BMIGRP_C6_V3=3;
Else if 30<= BMI_V3 < 35 then BMIGRP_C6_V3=4;
Else if 35<= BMI_V3< 40 then BMIGRP_C6_V3=5;
Else if BMI_V3>= 40 then BMIGRP_C6_V3=6;

Response format: 1=Underweight (BMI < 18.5)
                    2=Normal (18.5 ≤ BMI < 25)
                    3=Overweight (25 ≤ BMI < 30)

4=Obese I (30 ≤ BMI < 35)
5=Obese II (35 ≤ BMI < 40)
6=Obese III

Source variable(s):
BMI_V3. Body Mass Index (Kg/m^2) at Visit 3

7.4   CENTROBESE_V3 (Central Obesity at Visit 3)

Central Obesity is defined as a waist circumference (WC) ≥102 cm in assigned male at birth and ≥88 cm in assigned female at birth.

If (SEX_V3 = 0 and ANT10a >= 88) or (SEX_V3 = 1 and ANT10a >= 102) then CENTROBESE_V3 = 1;
Else CENTROBESE_V3 = 0;

Source variable(s):
SEX_V3. Sex assigned at birth updated at Visit 3
ANT10a. Waist girth measured in centimeters

7.5   WAIST_HIP_V3 (Waist to hip ratio at Visit 3)

This is a numerical variable providing the waist to hip ratio. This is obtained by simply dividing the waist girth by the hip girth. If either of these two values is missing, then the ratio is considered to be missing.

If neither waist girth nor hip girth are missing then WAIST_HIP_V3=ANT10A/ANT10B.

Source variable(s):
ANT10A. Waist girth measured in centimeters
ANT10B. Hip girth measured in centimeters

8.   BLOOD PRESSURE MEASURES

8.1   HYPERTMED_SELF_V3 (Hypertension – self-reported med use, Visit 3)

This is a 0/1 numeric variable. This variable combines self-reported antihypertensive medication use along with any indication of taking meds in last 4 weeks. Participants who had missing data for medication use were set to 0 (no medication use) when confirmed not taking any medications; participants who were unsure for medication use were set to missing.

If MUE26d=0 then HYPERTMED_SELF_V3=0;
Else if MUE26d=1 then HYPERTMED_SELF_V3=1;

Response format: 0=Self-reported medication use
                    1=No self-reported medication use

Source Variable(s):
MUE2. Is this because you forgot, because you have not taken any medications at all in the last four weeks, or because you could not bring your medications?
MUE26d. Self-reported high blood pressure or hypertension medication use
Note: MUE26d = missing had been reset to 0 when MUE2=1 (took no medication); MUE26d = 9 (unknown) had been reset to missing.

8.2   HYPERTENSION2_V3 (Hypertension – NHANES, Visit 3)

This is a 0/1 numeric variable. Hypertension is defined here using the following NHANES definition: if the systolic BP is greater than or equal to 140, or if the diastolic BP is greater than or equal to 90, or if the participant self-reports antihypertensive medication use in last 4 weeks.

if SBP5 >= 140 or SBP6 >= 90 or HYPERTMED_SELF_V3 = 1 then HYPERTENSION2_V3 = 1;
Else HYPERTENSION2_V3 = 0;
If all source variables are missing then HYPERTENSION2_V3 is missing;

Response format: 0=Not hypertensive (No)
                    1=Hypertensive (Yes)

Source variable(s):
SBP5. Systolic blood pressure
SBP6. Diastolic blood pressure
HYPERTMED_SELF_V3. Hypertension Med Use – self report only at Visit 3

8.3   HYPERTENSION2_C4_V3 (4-level grouped hypertension – NHANES, Visit 3)

Classifies participants into 4 categories based on hypertension using NHANES definition and self-reported antihypertensive medication use in last 4 weeks: no hypertension, pre-hypertension, treated hypertension, and untreated hypertension.

If HYPERTENSION2_V3 = 0 and (120<=SBP5<=140 or 80<=SBP6<=90) → 2
Else if HYPERTENSION2_V3 = 0 → 1
Else if HYPERTENSION2_V3 = 1 and HYPERTMED_SELF_V3 = 1 → 3
Else if HYPERTENSION2_V3 = 1 → 4

Response format: 1=Not hypertension
                    2=Pre-hypertension
                    3=Treated hypertension
                    4=Untreated hypertension

Source variable(s):
HYPERTENSION2_V3. Hypertension – NHANES, Visit 3
SBP5. Systolic blood pressure
SBP6. Diastolic blood pressure
HYPERTMED_SELF_V3. Hypertension Med Use – self report only at Visit 3

8.4   HYPERTENSION2_AHA_V3 (Hypertension – ACC/AHA, Visit 3)

This is a 0/1 numeric variable. Hypertension is defined here using the following ACC/AHA definition: if the systolic BP is greater than or equal to 130, or if the diastolic BP is greater than or equal to 80, or if the participant self-reports antihypertensive medication use in last 4 weeks.

If SBP5 >= 130 or SBP6 >= 80 or HYPERTMED_SELF_V3 = 1 then
HYPERTENSION2_AHA_V3 = 1;
Else HYPERTENSION2_AHA_V3 = 0;
If all source variables are missing then HYPERTENSION2_AHA_V3 is missing;

Response format: 0=Not hypertensive (No)
                    1=Hypertensive (Yes)

Source variable(s):
SBP5. Systolic blood pressure
SBP6. Diastolic blood pressure
HYPERTMED_SELF_V3. Hypertension Med Use – self report only at Visit 3

8.5   HYPERTENSION2_AHA_C5_V3 (5-level grouped hypertension – ACC/AHA, Visit 3)

Classifies participants into 5 categories based on hypertension using ACC/AHA definition and self-reported antihypertensive medication use in last 4 weeks: no hypertension, elevated hypertension, treated hypertension, untreated stage-1 hypertension, and untreated stage-2 hypertension.

If HYPERTENSION2_AHA_V3=0 and (120<=SBP5<130 and SBP6<80) $\rightarrow$ 2
Else if HYPERTENSION2_AHA_V3=0 $\rightarrow$ 1
Else if HYPERTENSION2_AHA_V3=1 and HYPERTMED_SELF_V3 =1 $\rightarrow$ 3
Else if HYPERTENSION2_AHA_V3=1 and (SBP5>=140 or SBP6>=90) $\rightarrow$ 5
Else if HYPERTENSION2_AHA_V3=1 $\rightarrow$ 4

Response format: 1=No hypertension
                    2=Elevated hypertension
                    3=Treated hypertension
                    4=Untreated Stage-1 hypertension
                    5=Untreated Stage-2 hypertension

Source variable(s):
HYPERTENSION2_AHA_V3. Hypertension – ACC/AHA, Visit 3
SBP5. Systolic blood pressure
SBP6. Diastolic blood pressure
HYPERTMED_SELF_V3. Hypertension Med Use – self report only at Visit 3

8.6   HYPER_DIAG_V3 (Hypertension – self-reported diagnosis, Visit 1, 2, or 3)

This is a 0/1 numeric variable that indicates if a participant has ever self-reported hypertension (from a doctor's diagnosis) at Visit 1, Visit 2, or Visit 3.

If baseline MHEA1=1 or V2 MHE15=1 or V3 MHE17=1 then HYPER_DIAG_V3 =1;
Else HYPER_DIAG_V3 = 0;
If all source variables are missing then HYPER_DIAG_V3 is missing;

Response format: 0=No
                                1=Yes

Source variable(s):
MHEA1. Has a doctor ever said that you have high blood pressure or hypertension? (V1)
MHE15. Has a doctor ever said that you have high blood pressure or hypertension? (V2)
MHE17. Has a doctor ever said that you have high blood pressure or hypertension? (V3)

9.   LABORATORY MEASURES

9.1   DYSLIPIDMED_SELF_V3 (Dyslipidemia – self-reported med use, Visit 3)

This is a 0/1 numeric variable. This variable combines self-reported dyslipidemia medication use along with any indication of taking meds in last 4 weeks. Participants who had missing data for medication use were set to 0 (no medication use) when confirmed not taking any medications; participants who were unsure for medication use were set to missing.

If MUE26e = 0 then DYSLIPIDMED_SELF_V3 = 0;
Else if MUE26e = 1 then DYSLIPIDMED_SELF_V3 = 1;

Source variable(s):
MUE2. Took no medication
MUE26e. Self-reported high blood cholesterol medication
intake

9.2   DYSLIPIDEMIA_V3 (Dyslipidemia/high cholesterol at Visit 3)

Dyslipidemia/High Cholesterol is defined as triglycerides>=200 or HDL<40 or LDL>=160. Additionally, if the participant self-reported taking anti-dyslipidemics they are coded as having Dyslipidemia/High Cholesterol.

If LABA107 >=200 or LABA68 < 40 or LABA69 >= 160 or DYSLIPIDMED_SELF_V3 = 1 then DYSLIPIDEMIA_V3 = 1;
Else DYSLIPIDEMIA_V3 = 0;
If LABA107 and LABA68 and LABA69 and DYSLIPIDMED_SELF_V3 are missing then DYSLIPIDEMIA_V3 = .;

Source variable(s):

LABA68. HDL-cholesterol (mg/dL)
LABA69. LDL-cholesterol (mg/dL)
LABA107. Triglycerides (mg/dL)
DYSLIPIDMED_SELF_V3. Dyslipidemia Med Use – self report only at Visit 3

9.3 DYSLIPIDEMIA_C3_V3 (3-Level Grouped Dyslipidemia, Visit 3)

This is a 3-category variable that defines whether a participant has dyslipidemia, and if so, what the treatment status of the condition is.

If DYSLIPIDEMIA_V3 or MUE26E are missing then DYSLIPIDEMIA_C3_V3=.;
Else if DYSLIPIDEMIA_V3=0 then DYSLIPIDEMIA_C3_V3=1;
Else if DYSLIPIDEMIA_V3=1 & MUE26E=0 then DYSLIPIDEMIA_C3_V3=2;
Else if DYSLIPIDEMIA_V3=1 & MUE26E=1 then DYSLIPIDEMIA_C3_V3=3;

Response format: 1=Not dyslipidemic (No)
                    2=Dyslipidemic without treatment
                    3=Dyslipidemic undergoing treatment

Source variable(s):
DYSLIPIDEMIA_V3. Dyslipidemia status at Visit 3
MUE26E. Self-reported high blood cholesterol medication intake
Note: MUE26E missing had been reset to 0 when MUE2=1 (confirmed not taking any medications).

9.4 DYSLIPIDEMIA_TCHDL_V3 (Dyslipidemia based on TCHDL>5 at Visit 3)

This is a 0/1 variable that checks for dyslipidemia as measured by a high total cholesterol to HDL cholesterol ratio.

If CARDIAC_RISK_RATIO_V3 <=.Z then DYSLIPIDEMIA_TCHDL_V3=.;
Else if CARDIAC_RISK_RATIO_V3 >5 then DYSLIPIDEMIA_TCHDL_V3=1;
Else DYSLIPIDEMIA_TCHDL_V3=0;

Response format: 0=Not dyslipidemic (No)
                    1=Dyslipidemic (Yes)

Source variable(s):
CARDIAC_RISK_RATIO_V3. Cardiac Risk Ratio (TC/HDL) – AHA at Visit 3

9.5 DYS_TCHDL_MED_V3 (Dyslipidemia based on TCHDL>5 or lipid reduction drug use at Visit 3)

This is a 0/1 variable that checks for dyslipidemia as measured by either a high total cholesterol to HDL ratio, or by self-reported high blood cholesterol medication intake.

If DYSLIPIDEMIA_TCHDL_V3 & MUE26E both missing then DYS_TCHDL_MED_V3=.;

Else if DYSLIPIDEMIA_TCHDL_V3=1 or MUE26E = 1 then DYS_TCHDL_MED_V3=1;
Else DYS_TCHDL_MED_V3=0;

Response format: 0=Not dyslipidemic (No)
                    1=Dyslipidemic (Yes)

Source variable(s):
DYSLIPIDEMIA_TCHDL_V3. Dyslipidemia based on TCHDL>5 at Visit 3
MUE26E. Self-reported high blood cholesterol medication intake
Note: MUE26E missing had been reset to 0 when MUE2=1 (confirmed not taking any medications).

9.6   DYSLIPID_DIAG_V3 (Dyslipidemia/High Cholesterol – self-reported diagnosis, Visit 1, 2, or 3)

This is a 0/1 numeric variable that indicates if a participant has ever self-reported dyslipidemia (from a doctor's diagnosis) at Visit 1, Visit 2, or Visit 3.

If baseline MHEA2=1 or V2 MHE16=1 or V3 MHE18=1 then DYSLIPID_DIAG_V3 =1;
Else DYSLIPID_DIAG_V3 = 0;
If all source variables are missing then DYSLIPID_DIAG_V3 is missing;

Response format: 0=No
                    1=Yes

Source variable(s):
MHEA2. Has a doctor ever said that you have high blood cholesterol? (V1)
MHE16. Has a doctor ever said that you have high blood cholesterol? (V2)
MHE18. Has a doctor ever said that you have high blood cholesterol? (V3)

9.7   CARDIAC_RISK_RATIO_V3 (Cardiac Risk Ratio (TC/HDL) – AHA at Visit 3)

This is a numeric variable that calculates the ratio of total cholesterol by HDL cholesterol, as follows:

If LABA66 and LABA68 are both non-missing, then:
CARDIAC_RISK_RATIO_V3=LABA66/LABA68

Source variable(s):
LABA66. Total Cholesterol (mg/dL)
LABA68. HDL - cholesterol (mg/dL)

9.8   TOTAL_CHOL_C3_V3 (Total Cholesterol at Visit 3)

This is a 3-category variable defining a participant's cholesterol status.

   If .Z < LABA66 < 200 then TOTAL_CHOL_C3_V3=1;
   Else if 200 <= LABA66 <= 239 then TOTAL_CHOL_C3_V3=2;

Else if LABA66 >=240 then TOTAL_CHOL_C3_V3=3;
Else TOTAL_CHOL_C3_V3=.;

Response format: 1=Normal Cholesterol
                    2=Borderline High Cholesterol
                    3=High Cholesterol

Source variable(s):
LABA66. Total Cholesterol (mg/dL)

9.9 HSCRP_AVAIL_V3 (High-sensitivity C-Reactive Protein Available (SGM SOL ancillary study))

This variable indicates if a participant has the continuous measure of high-sensitivity C-reactive protein (hsCRP). HSCRP was only measured in Visit 3 as part of the SGM SOL ancillary study.

Response format: 0=Does not have hsCRP result
                    1=Has hsCRP result

9.10 FASTTIME_V3 (Fasting time (hrs.))

This variable combines fasting time variables collected from the Biospecimen Collection (BIO) form, as follows (for non-missing BIO4):

If BIO4 > .Z then do;
    If (BIO4=1) | (BIO5='0:00't) then t=0;
    Else if BIO4=2 then t=1;
    Else if BIO4=3 then t=2;
    FASTTIME_V3=((((BIO7+(t*'24:00't))-BIO5)/60)/60);
End;

Source variable(s):
BIO4. On which day did you last eat or drink anything except water: today, yesterday, or the day before yesterday?
BIO5. And at what time was that?
BIO7. Collection time

10. CLINICAL CHARACTERISTICS

10.1 DIABETES_LAB_V3 (3-level grouped diabetes – ADA lab, Visit 3)

This is a numeric variable that describes the stages of diabetes based on American Diabetes Association (ADA) lab criteria. This definition takes into account serum glucose levels adjusted for fasting time, post-OGTT glucose levels if available, and A1C percentages. This variable does NOT include any medication use variables.

If (fasting time > 8 hrs and fasting glucose >= 126 mg/dL) or (fasting time <= 8 and fasting glucose >= 200 mg/dL) or (post-OGTT glucose >= 200 mg/dL) or (A1C >= 6.5%) → 3
Else if (fasting time > 8 hrs and 100 mg/dL <= fasting glucose <= 125 mg/dL) or (140 mg/dL <= post-OGTT glucose <= 199 mg/dL) or (5.7% <= A1C < 6.5%) then → 2
Else → 1
If fasting glucose, post-OGTT, and A1C are missing → missing

Response format: 1=Normal glucose regulation
                    2=Impaired glucose tolerance
                    3=Diabetes

Source variable(s):
LABA70. Fasting glucose (mg/dL)
LABA72. % Glycosylated Hemoglobin (A1C)
FASTTIME_V3. Elapsed time (hrs) between the time the participant last consumed anything and the blood draw
Note: Post-OGTT glucose level was not measured at Visit 3

10.2   SELFMED_ANTIDIAB_V3 (Diabetes – self-reported med use, Visit 3)

This is a 0/1 numeric variable. This variable combines self-reported antidiabetics medication use along with any indication of taking meds in last 4 weeks. Participants who had missing data for medication use were set to 0 (no medication use) when confirmed not taking any medications; participants who were unsure for medication use were set to missing.

If MUE26c=0 then SELFMED_ANTIDIAB_V3=0;
Else if MUE26c=1 then SELFMED_ANTIDIAB_V3=1;

Response format: 0=Self-reported medication use
                    1=No self-reported medication use

Source variable(s):
MUE2. Is this because you forgot, because you have not taken any medications at all in the last four weeks, or because you could not bring your medications?
MUE26c. Self-reported high blood sugar or diabetes medication use
Note: MUE26c = missing had been reset to 0 when MUE2=1 (took no medication); MUE26c = 9 (unknown) had been reset to missing.

10.3   DIABETES4A_V3 (3-level grouped diabetes – def. 4A (ADA lab no OGTT & self-reported med use), Visit 3)

This is a numeric variable that describes the stages of diabetes based on ADA lab criteria (taking into account serum glucose levels adjusted for fasting time, post-OGTT glucose levels if available, and A1C percentages) and self-reported antidiabetics medication use. Since post-OGTT glucose levels are not available at Visit 3, the letter 'A' is added to the Visit 3 definition name to highlight this distinction compared to previous visits. This

diabetes variable was defined by the algorithm strictly. Per Diabetes SIG decision and memo, Diabetes classifications at previous visits are not part of definition specs and the HCHS/SOL study lab classification was not considered as a formal diagnosis.

If SELFMED_ANTIDIAB_V3=1 or DIABETES_LAB_V3=3 then DIABETES4A_V3=3;
Else if DIABETES_LAB_V3=2 then DIABETES4A_V3=2;
Else DIABETES4A_V3=1;
If all source variables are missing then DIABETES4A_V3 is missing;

Response format: 1=Non-diabetic
                    2=Pre-diabetic
                    3=Diabetic

Source variable(s):
SELFMED_ANTIDIAB_V3. Diabetes – self-reported med use, Visit 3
DIABETES_LAB_V3. 3-level grouped diabetes – ADA lab, Visit 3

10.4   DIABETES4A_C4_V3 (4-level grouped diabetes – def. 4A (ADA lab no OGTT & self-reported med use), Visit 3)

This is a numeric variable that describes the stages of diabetes with/without treatment based on ADA lab criteria (taking into account serum glucose levels adjusted for fasting time, post-OGTT glucose levels if available, and A1C percentages) and self-reported antidiabetics medication use. Since post-OGTT glucose levels are not available at Visit 3, the letter 'A' is added to the Visit 3 definition name to highlight this distinction compared to previous visits.

DIABETES4A_V3=1 → 1
DIABETES4A_V3=2 → 2
DIABETES4A_V3=3 and (SELFMED_ANTIDIAB_V3 is 1) → 3
DIABETES4A_V3=3 and (SELFMED_ANTIDIAB_V3 is not 1) → 4
Missing DIABETES4A_V3 → missing

Response format: 1=Non-diabetic
                    2=Pre-diabetic
                    3=Diabetic with treatment
                    4=Diabetic without treatment

Source variable(s):
DIABETES4A_V3. 3-level grouped diabetes – def. 4A (ADA lab no OGTT & self-reported med use), V3
SELFMED_ANTIDIAB_V3. Diabetes – self-reported med use, Visit 3

10.5 DIABETES4A_INDICATOR_V3 (Indicator of diabetes – def. 4A (ADA lab no OGTT & self-reported med use), Visit 3)

This is a 0/1 indicator variable that describes the classification of diabetes based on ADA lab criteria (taking into account serum glucose levels adjusted for fasting time, post-OGTT glucose levels if available, and A1C percentages) and self-reported antidiabetics medication use. Since post-OGTT glucose levels are not available at Visit 3, the letter 'A' is added to the Visit 3 definition name to highlight this distinction compared to previous visits. This indicator variable could be used as an incident diabetes case indicator within the "not diabetic" group classified by diabetes definition 4 at a previous visit.

If DIABETES4A_V3=3 then DIABETES4A_INDICATOR_V3=1;
Else if DIABETES4A_V3 in (1,2) then DIABETES4A_INDICATOR_V3=0;
Else DIABETES4A_INDICATOR_V3=.;

Response format: 0=Not diabetic (No)
                    1=Diabetes (Yes)

Source variable(s):
DIABETES4A_V3. 3-level grouped diabetes – def. 4 (ADA lab & self-reported med use), V3

10.6 DIABETIC_DIAG_V3 (Diabetes – self-reported diagnosis, Visit 1, 2, or 3)

This is a 0/1 numeric variable that indicates if a participant has ever self-reported diabetes (from a doctor's diagnosis) at Visit 1, Visit 2, or Visit 3. It does not exclude gestational diabetes.

If V1 MHEA16=1 or V2 MHE14=1 or V3 MHE16=1 then DIABETIC_DIAG_V3 =1;
Else DIABETIC_DIAG_V3 = 0;
If all source variables are missing then DIABETIC_DIAG_V3 is missing;

Response format: 0=No
                    1=Yes
Source variable(s):
MHEA16. Has a doctor ever said that you have diabetes (high sugar in blood or urine)? (V1)
MHE14. Has a doctor ever said that you have diabetes (high sugar in blood or urine)? (V2)
MHE16. Has a doctor ever said that you have diabetes (high sugar in blood or urine)? (V3)

10.7 METS_NCEP2_V3 (Metabolic Syndrome w/ self-reported meds at Visit 3 - NCEP-ATP III)

This is a 0/1 numeric variable with 1=metabolic syndrome present and 0=no metabolic syndrome in Visit 3. This variable incorporates the original METS_NCEP definition with changes to include self-reported medication use instead of scanned/transcribed medication use.

The diagnosis of metabolic syndrome is made when three or more of the following risk factors are present:

Elevated_BP_SelfMeds: SBP5 ≥ 130 or SBP6 ≥ 85 mm Hg or self-report use of antihypertension medications (HYPERTMED_SELF_V3=1) (note: requires non-missing SPB, DBP, or self-report medication use)
High_Trig: LABA107 ≥ 150 mg/dL
Low_HDL: LABA68 < 40 mg/dl (assigned male sex at birth), LABA68 < 50 mg/dl (assigned female sex at birth)
IFG_NCEP_SelfMeds: LABA70 ≥ 100 mg/dl or self-report use of anti-diabetic medications (SELFMED_ANTIAB_V3=1)
Abdominal_obesity_NCEP: CENTROBESE_V3 = 1

First, each of the components is evaluated (0/1, or missing). If 3 or more components are missing components, then METS_NCEP2_V3 should be missing.

Otherwise, METS_NCEP2_V3 = 0 if sum of five components is 0, 1, 2
= 1 if sum of five components is 3, 4, 5

Source Variable(s):
SBP5. Systolic blood pressure
SBP6. Diastolic blood pressure
HYPERTMED_SELF_V3. Hypertension Med Use – self report only at Visit 3
LABA68. HDL-cholesterol (mg/dL)
SEX_V3. Sex assigned at birth updated at Visit 3
LABA70. Fasting glucose (mg/dL)
LABA107. Triglycerides (mg/dL)
SELFMED_ANTIDIAB_V3. Self-report or medication list report taking anti-diabetic medication
CENTROBESE_V3. Central obesity at Visit 3

10.8   METS_V3 (Metabolic Syndrome at Visit 3)

This is a 0/1 numeric variable with 1=metabolic syndrome present and 0=no metabolic syndrome in Visit 3. This variable incorporates self-reported dyslipidemia medication use.

The diagnosis of metabolic syndrome is made when three or more of the following risk factors are present:

Elevated_BP: SBP5 ≥ 130 or SBP6 ≥ 85 mm Hg
SelfMeds: Self-report use of antihypertension medications (HYPERTMED_SELF_V3=1), dyslipidemia medications (DYSLIPIDMED_SELFV3), or anti-diabetic medications (SELFMED_ANTIAB_V3=1)
High_Trig: LABA107 ≥ 150 mg/dL
Low_HDL: LABA68 < 40 mg/dl (assigned male sex at birth), LABA68 < 50 mg/dl (assigned female sex at birth)
IFG_NCEP: LABA70 ≥ 100 mg/dl or
Abdominal_obesity_NCEP: CENTROBESE_V3 = 1

First, each of the components is evaluated (0/1, or missing). If all components are missing components, then METS_V3 is missing.

Otherwise, METS_V3   = 0 if sum of five components is < 3
                         = 1 if sum of five components is ≥ 3

Source Variable(s):
SBP5. Systolic blood pressure
SBP6. Diastolic blood pressure
HYPERTMED_SELF_V3. Hypertension Med Use – self report only at Visit 3
DYSLIPIDMED_SELFV3. Dyslipidemia Med Use – self report only at Visit 3
SELFMED_ANTIDIAB_V3. Self-report or medication list report taking anti-diabetic medication
SEX_V3. Sex assigned at birth updated at Visit 3
LABA68. HDL-cholesterol (mg/dL)
LABA70. Fasting glucose (mg/dL)
LABA107. Triglycerides (mg/dL)
CENTROBESE_V3. Central obesity at Visit 3

10.9   CURRENT_SMOKER_V3 (Current smoker status at Visit 3)

This is a 0/1 indicator variable to denote current smoker status at Visit 3 using the variables TBE1 and TBE2.

  If TBE1=1 & TBE2 in (1,2) then CURRENT_SMOKER_V3=1;
  Else if (TBE1=1 & TBE2 = 3) or TBE1=0 then CURRENT_SMOKER_V3=0;
Response format: 0 = No (i.e., never or former smoker)
                    1 = Yes (i.e., current smoker)

Source variable(s):
TBE1. Have you ever smoked at least 100 cigarettes in your entire life?
TBE2. Do you NOW smoke daily, some days or not at all?

10.10   HBA1C_SI_V3 (Respondent's HbA1c levels (in mmol/mol))

HBA1C_SI_V3 = (LABA72 *10.93) – 23.50, rounded to the nearest integer

Beginning with manuscripts submitted after Jan. 1, 2013, Diabetes requires authors to report HbA1c levels in both traditional, DCCT-derived units (as %) and SI, IFCC-recommended units (as mmol/mol).

Reference: Change of HbA1c reporting to the new SI units, Jones G, Barker G, et al. MJA 2011; 195: 45–46

Source variable(s):
LABA72. Glycosylated Hemoglobin (%)

10.11   HIGH_TOTAL_CHOL2_V3 (Hypercholesterolemia at Visit 3 (from LABA66, LABA68, LABA69, MUE26E))

This is a 0/1 variable that checks for Total cholesterol, HDL, LDL, and self-reported antihyperlipidemic medication use values to determine presence/absence of hypercholesterolemia as follows:

If missing(LABA66) & missing(LABA68) & missing(LABA69) & (missing(MUE26E or MUE26E=9) then HIGH_TOTAL_CHOL2_V3=missing
Else if (LABA66>=240) OR (LABA69>=160) OR (0<=LABA68<40) OR (MUE26E=1) then HIGH_TOTAL_CHOL2_V3=1;
Else HIGH_TOTAL_CHOL2_V3=0;

Response format: 0=Not Hypercholesterolemic (No)
                    1=Hypercholesterolemic (Yes)

Source variable(s):
LABA66. Total Cholesterol (mg/dL)
LABA68. HDL - Cholesterol (mg/dL)
LABA69. LDL - Cholesterol (mg/dL)
MUE26E. Self-reported use lipid lowering drugs/antihyperlipidemic medication.
Note: MUE26E missing had been reset to 0 when MUE2=1 (confirmed not taking any medications).

Variable Modifications

Date release	Variable	Number of records (participants) changed	Notes
09/2024	EDUCATION_C3_V3	989	Algorithm corrected to use EDUCATION_C3 from baseline if missing EDUCATION_C3_V2.
	WEIGHT_EXPANDED_ALL_V3		NEW VARIABLE
	WEIGHT_NORM_OVERALL_ALL_V3		NEW VARIABLE
	WEIGHT_NORM_CENTER_ALL_V3		NEW VARIABLE
	WEIGHT_EXPANDED_EXAMONLY_V3		NEW VARIABLE
	WEIGHT_NORM_OVERALL_EXAMONLY_V3		NEW VARIABLE
	WEIGHT_NORM_CENTER_EXAMONLY_V3		NEW VARIABLE
	PARTICIPANT_EXAMONLY_V3		NEW VARIABLE
	PHONE_EXAM_CONVERT_V3		NEW VARIABLE
	VISIT_TYPE_V3		NEW VARIABLE
	HOMEVISIT_V3		NEW VARIABLE
	ANY_PERMIT_V3		NEW VARIABLE
	DURATION_V3		NEW VARIABLE
	AGEGRP_C7_V3		NEW VARIABLE
	SEX_C3_V3		NEW VARIABLE
	INTERSEX		NEW VARIABLE
	SGM		NEW VARIABLE
	TRANSGENDER		NEW VARIABLE
	GENIDENTITY_C7		NEW VARIABLE
	GENIDENTITY_C4		NEW VARIABLE
	CISGENDER		NEW VARIABLE
	SXORIENTATION		NEW VARIABLE
	TNB		NEW VARIABLE
	LGBTID		NEW VARIABLE
	SAMESXATT		NEW VARIABLE
	SAMESXPART		NEW VARIABLE
	OPPSXPART		NEW VARIABLE
	ASEXUAL		NEW VARIABLE
	TGD		NEW VARIABLE
	AGE_GENDERID_LE_AGE_V1		NEW VARIABLE
	AGE_SOID_LE_AGE_V1		NEW VARIABLE
	INCOME2_V3		NEW VARIABLE
	INCOME2_C3_V3		NEW VARIABLE
	INCOME2_C5_V3		NEW VARIABLE
	EMPLOYEDFY_V3		NEW VARIABLE
	MARITAL_STATUS_V3		NEW VARIABLE
	HHSIZE_V3		NEW VARIABLE
	HC_V3		NEW VARIABLE
	LANG_PREF_V3		NEW VARIABLE
	YRSUS2_C2_V3		NEW VARIABLE
	YRSUS2_C3_V3		NEW VARIABLE
	DISCRIMEVER_V1V3		NEW VARIABLE
	IMMSTRESSSUM_V3		NEW VARIABLE
	IMMSTRESSAVG_V3		NEW VARIABLE
	CHRSTRTOT_V3		NEW VARIABLE
	CHRSTRTOT6_V3		NEW VARIABLE

	CHRSTRSEV_V3		NEW VARIABLE
	M_V_TOT_V3		NEW VARIABLE
	BRS_V3		NEW VARIABLE
	BRSAVG_V3		NEW VARIABLE
	PHQ8_V3		NEW VARIABLE
	PHQ8_GE10_V3		NEW VARIABLE
	BMI_V3		NEW VARIABLE
	BMI_C4_V3		NEW VARIABLE
	BMI_C6_V3		NEW VARIABLE
	CENTROBESE_V3		NEW VARIABLE
	WAIST_HIP_V3		NEW VARIABLE
	HYPERTMED_SELF_V3		NEW VARIABLE
	HYPERTENSION2_V3		NEW VARIABLE
	HYPERTENSION2_C4_V3		NEW VARIABLE
	HYPERTENSION2_AHA_V3		NEW VARIABLE
	HYPERTENSION2_AHA_C5_V3		NEW VARIABLE
	HYPER_DIAG_V3		NEW VARIABLE
	DYSLIPIDMED_SELF_V3		NEW VARIABLE
	DYSLIPIDEMIA_V3		NEW VARIABLE
	DYSLIPIDEMIA_C3_V3		NEW VARIABLE
	DYSLIPIDEMIA_TCHDL_V3		NEW VARIABLE
	DYS_TCHDL_MED_V3		NEW VARIABLE
	DYSLIPID_DIAG_V3		NEW VARIABLE
	CARDIAC_RISK_RATIO_V3		NEW VARIABLE
	TOTAL_CHOL_C3_V3		NEW VARIABLE
	FASTTIME_V3		NEW VARIABLE
	DIABETES_LAB_V3		NEW VARIABLE
	SELFMED_ANTIDIAB_V3		NEW VARIABLE
	DIABETES4A_V3		NEW VARIABLE
	DIABETES4A_C4_V3		NEW VARIABLE
	DIABETES4A_INDICATOR_V3		NEW VARIABLE
	DIABETIC_DIAG_V3		NEW VARIABLE
	METS_NCEP2_V3		NEW VARIABLE
	METS_V3		NEW VARIABLE
	CURRENT_SMOKER_V3		NEW VARIABLE
	HBA1C_SI_V3		NEW VARIABLE
	HSCRP_AVAIL_V3		NEW VARIABLE
	HIGH_TOTAL_CHOL2_V3		NEW VARIABLE