NHLBI grant: Ancillary to HCHS/SOL: Hispanic Community Children's Health Study of Latino Youth (R01 HL102130; S Bangdiwala, G Ayala, M Carnethon, A Delamater, J Eckfeldt, J Himes, C Isasi, & K Perreira)

Hispanic Community Children's Health/ Study of Latino Youth (SOL Youth)

Derived Variable Dictionary - Child

September 2018 - Version INV2.2

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

Krista M. Perreira
Franklyn Gonzalez II

HCHS/SOL Youth Derived Variable Dictionary
Table of Contents

Version History

Version	Who	Date	What changed
INV1	Franklyn	September 2014	Original
INV1_1	Franklyn	May 2015	Changes: PREDIABETES updated to reflect HbA1c cutoff change to 5.7 (from 6.0) Additions: PAE_TRANSPORT, PAE_SCHOOL, PAE_LEISURENONSPORT, PAE_SPORT, PAE_HOUSEHOLD, PAE_SCREEN_MIN, PAE_SED, PAE_LIGHT, PAE_MOD, PAE_VIG, PAE_MVPA, SBP_FLAG, DBP_FLAG, HBA1C_FLAG, HOMA_IR, INSULIN_RESIST, OBESITY_BIN, CVD_RISK_COUNT, CVD_RISK_C4, CVD_RISK_C3, CVD_RISK_BIN, YOUTH_WEIGHT_CALIB, YOUTH_WEIGHT_NORM_OVERALL, YOUTH_WEIGHT_NORM_CENTER
INV1_2	Franklyn	July 2015	Update HOMA_IR definition (and therefore INSULIN_RESIST) to calibrate insulin variable LABA11 units from pmol/L to mU/mL in the formula (divide by 6).
INV2_0	Franklyn	February 2016	Additions: VO2_MAX (Estimated Maximum VO2 – fitness step test) HOUSEHOLD_LINK (Household identifier of child participant – masked) HEI-2010 score and its 12 components scores plus intermediate variables used to create it. Updates: DIABETES_GLUCOSE – Updated cutoff to be >= 126 instead or >126 in previous version. Note that no participant was affected by this change.
INV2_1	Franklyn	May 2016	Additions: AGEMOS - Age In Months At The Time Of Child's Clinic Visit BMIPCT95 - Body Mass Index 95th Percentile -CDC (Child) Updates: STRAT – Low sample size STRAT values have been collapsed into another STRAT group with similar sampling qualities (ex. Hispanic household proportion) within each field center. FPG_FLAG – Updated cutoff to be >= 100 instead or >100 in previous version. HBA1C_FLAG – Updated cutoff to be >= 5.7 instead or >5.7 in previous version. INSULIN_RESIST - Updated cutoff to be >= 2.5 instead or >2.5 in previous version. Note that no participant was affected by this change. WHOLE_GRAINS_R1 and WHOLE_GRAINS_R2: Updated to correct coding typo by replacing Y09A41with Y09A43. Variables affected by FPG_FLAG and/or HBA1C_FLAG definition changes: MET_SYNDROME_ATP, MET_SYNDROME_IDF, CVD_RISK_COUNT, CVD_RISK_C4, CVD_RISK_C3, CVD_RISK_BIN

			Variables affected by WHOLE_GRAINS_R1 and/or WHOLE_GRAINS_R2 definition changes: WHOLE_GRAINS_AVG, WHOLE_GRAINS_DEN_R1, WHOLE_GRAINS_DEN_R2, WHOLE_GRAINS_DEN_AVG, HEI5, HEI2010
INV2_2	Franklyn	September 2018	Updates: BMIGRP_C5 – Correct typo in categories 4 and 5 to reflect 120% instead of 125% was used and clarify and/or statements. *****Note: Algorithm definition is correct, 125% was only a typo in the data dictionary text.**

1. DESIGN

1.1 ID: SOL Youth Participant ID - Masked
Indicator of unique participant ID that has been masked to prevent de-identification.

1.2 ID_TYPE: Child Or Parent ID - Masked
Study variable denoting an ID that reflects a study child participant or a parent/LG.

Response format: 1 = Child study participant
2 = Parent/LG study participant

1.3 CHILD_PARENT_LINK: Parent Id Linked To Child
Study variable denoting a parent/LG participant ID linked to each child.

1.4 HOUSEHOLD_LINK: Household identifier of child participant - masked
Study variable denoting a household identifier linked to each child. Children coming from the same household will share the same HH_ID value.

1.5 PSU_ID: Primary Sampling Unit ID Clustering Variable- masked
HCHS/SOL Sample design clustering variable used in statistical analyses as the CLUSTER variable (in SAS) that is a combination of the following:

Field Center and selected block group identifier.

1.6 STRAT: Sampling Design Stratification Variable - masked
HCHS/SOL Sample design stratification variable used in statistical analyses as the STRATUM variable (in SAS) that is a combination of the following:

Field Center, Hispanic Household Proportion (high, low), and SES (high, low).

(Hispanic Household Proportion and SES calculated using 2000 Census Data for selected block groups within each field center)

1.7 YOUTH_WEIGHT_CALIB: SOL Youth Census 2010 Calibrated, Trimmed, Nonresponse Adjusted Weights - (frozen as of 4/2014)
The SOL Youth calibrated sample weight is the calibrated, trimmed, non-response adjusted reciprocal of a participant's probability of selection into the SOL Youth study. The trimmed, non-response adjusted sample weights were calibrated using first the 18-category age group/gender distribution within each field center using the 2010 Decennial Census. This variable is used for analyses purposes only.

1.8 YOUTH_WEIGHT_NORM_OVERALL: SOL Youth Overall Normalized, Census 2010 Calibrated, Trimmed, Nonresponse Adjusted Weights - (frozen as of 4/2014)

The SOL Youth normalized sample weight is the overall normalized, calibrated, trimmed, non-response adjusted (household and person-level) reciprocal of a participant's probability of selection into the SOL Youth study. The base sample weights, i.e. the reciprocal of a participant's probability of selection into the HCHS/SOL study, were first adjusted by household-level and person-level non-response in HCHS/SOL and then in SOL Youth. The non-response adjusted sample weights were then trimmed using the 95^{th} percentile in order to minimize MSE effects within each field center. All sample weights that fell above the selected percentile were trimmed to the selected percentile and the trimmed difference was evenly distributed among all of the non-trimmed sample weights. The trimmed, non-response adjusted sample weights were then calibrated using the 18-category age group/gender distribution using the 2010 Decennial Census. Finally, these calibrated, trimmed, non-response adjusted sample weights were normalized using the weighted mean divided by weighted standard deviation of the entire SOL Youth sample to normalize the expanded calibrated sample weights. This final sample weight variable is used for analyses purposes only.

1.9 YOUTH_WEIGHT_NORM_CENTER: 'SOL Youth Center Normalized, Census 2010 Calibrated, Trimmed, Nonresponse Adjusted Weights - (frozen as of 4/2014)

The SOL Youth center-specific normalized sample weight is the field center specific normalized, calibrated, trimmed, non-response adjusted (household and person-level) reciprocal of a participant's probability of selection into the SOL Youth study. The calibrated, trimmed, non-response adjusted sample weights were normalized using the weighted mean divided by weighted standard deviation calculated for each field center separately and used to normalize the expanded calibrated sample weights. This final sample weight variable is used for analyses purposes only.

2. ADMINISTRATIVE

2.1 CENTER: Participant's field center
This is a character variable with four possible values derived from the city of origin: "B"= Bronx, "C"= Chicago, "M"= Miami, "S"= San Diego. Center cannot have missing values because each valid participant ID has an affiliated field center.

Source variables: SUBJECTID

2.2 CENTERNUM: Numeric (Participant's field center)
This is a numeric variable with four possible values derived from the city of origin: 1= Bronx, 2= Chicago, 3= Miami, 4= San Diego. Center cannot have missing values because each valid participant ID has an affiliated field center.

Source variables: CENTER

2.3 CLINDATE: Date of the participant's clinic visit
This is a SAS date variable which documents the date of the participant's clinic visit. It is first derived from the Medical History (MHE). If the form or the date of form completion is not present, then it is derived from the Informed Consent tracking form (ICT), Anthropometry (ANT), or Sitting Blood Pressure (SBP) forms.

CLINDATE= MHE0A
or
Minimum (ICT0A, SBP0A, ANT0A) if MHE0A is missing

Source variable(s): ICT0A, MHE0A, SBP0A, ANT0A

2.4 CONSENT: Informed consent status of parent for participation in study
This is a binary variable that determines whether or not a parent study participant consented to participate in the baseline examination study (and did not withdraw consent).

If ICT1=1 and ICU1^=0 then consent=1;
Else consent=0;

Source variables: ICT1, ICU1

2.5 ASSENT: Informed Assent status of child/young adult for participation in study
This is a binary variable that determines whether or not a child/young adult study participant assented to participate in the baseline examination study (and did not withdraw assent).

If (IAT1=1 or IYT=1) and (IAU^=0 and IYU ^=0) then assent=1;
Else assent=0;

Source variables: IAT1, IYT1, IAU1, IYU1

2.6 PARTICIPATE: Informed Consent/Assent status for participation in study

This is a binary variable that determines whether or not both the parent and child/young adult study participants agreed to participate in the baseline examination study (and did not withdraw assent).

If CONSENT=1 and ASSENT=1 then PARTICIPATE=1;
Else PARTICIPATE=0;

Source variables: CONSENT, ASSENT

3. SOCIO-DEMOGRAPHIC

3.1 AGE: Age in years at the time of child's clinic visit

This is the age of the child participant in years (an integer variable) at the time of the participant's clinic visit. It is determined from the participant's date of birth and the clinic visit date.

AGE = INTEGER of (MHE1– MHE0A)/365.25

Source variables: MHE0A, MHE1

3.2 AGEMOS: Age In Months At The Time Of Child's Clinic Visit

This is the age of the child participant in months at the time of the participant's clinic visit. It is determined from the participant's date of birth and the clinic visit date.

AGEMOS = ((MHE1– MHE0A)/365.25)*12

Note: This variable can be used to calculate any additional age variables needed, as follows:
Age, in months, as integer value only: INTEGER of (AGEMOS).
Age, in years, as decimal value: AGEMOS/12
Age, in years, as integer value (equal to AGE): Integer of AGEMOS/12

Source variables: MHE0A, MHE1

3.3 AGEGROUP_C3: 3-level grouped age of child participant

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

AGE	AGEGROUP_C2
8-12	1
13-14	2
15-16	3

Response Format: 1=Ages 8-12
2=Ages 13-14
3=Ages 15-16

Source variables: AGE

3.4 GENDER: Gender of child participant

This is a categorical character variable which describes the child participant's gender, girl (G) or boy (B).

DCE1	Gender

1	B
2	G

Response Format: B=Boy
G=Girl

Source variables: DCE1

3.5 GENDERBIN: Gender – binary (0=Girl, 1=Boy))
This is a 0/1 variable which describes the child participant's gender, girl (0) or boy (1).

Gender	Gendernum
G	0
B	1

Response Format: 0=Girl
1=Boy

Source variables: GENDER

4. RACE AND ETHNIC IDENTIFICATION

4.1 RACE: Racial Self-Identification (Child)

Racial self-identifications are collapsed into categories with sufficient Ns for subgroup analyses. Category 3 includes children who identified as Hispanic/Latino (DCE2=1) or identified with a specific Hispanic/Latino group (DCE2a=0,1,2,3,4,5,6,7) but either reported their race as other, unknown, or not reported. Those remaining in the other category reported more than one race, Asian, or Native American.

Response format:
1 = Black
2 = White
3 = Hispanic/Latino Only
4 = Other

Source Variables: DCE3, DCE2, DCEa2

4.2 WHITE: Racial Self-Identification (Child)

This is a two-category racial self-identification distinguishing between those who identify themselves as white (RACE=2) and those who do not identify as white.

Response format:
1 = White
0 = Non-White

Source Variables: RACE

5. ACCULTURATION

5.1 USBORN_MAIN: Grouped Place of Birth (Child)

This variable follows the coding structure of the HCHS/SOL adult study and groups the place of birth of the participant to the United States (50 states only) or other place of birth. Other place of birth includes missing, Puerto Rico, and foreign countries.

Response format: 1 = US Mainland
0 = Not born in the US Mainland (includes Puerto Rico, missing)

Source Variables: DCE4

Note: This is the same as a two-category immigrant generation variable. Persons not born in the US mainland are first-generation immigrants to the US mainland. Persons born in the US mainland may be either second- or third+ generation immigrants. This follows common practice in demographic research of treating Puerto Rican born individuals as first-generation immigrants because their migration experiences can be similar.

Warning: This variable should not be used in analyses where health insurance, employment or college education are the primary outcomes. In these analyses, Puerto Ricans and others born in US territories should be treated as US born. Please use the USBORN_PR variable instead.

5.2 USBORN_PR: Grouped Place of Birth - Puerto Rican Revision (Child)

This variable follows the coding structure of the HCHS/SOL adult study and groups the place of birth of the participant as a US-born citizen (i.e. native) born in the mainland US or a US territory (e.g., Puerto Rico).

Response format: 1 = Born in US or US territory
0 = Not born in the US or US territory

Source Variables: DCE4

5.3 IMGEN: Immigrant Generation (Child)

This is a three category immigrant generation variable. First-generation children are defined as foreign-born with foreign-born parents. Second-generation children are defined as US-born with at least one foreign-born parent. Third-generation immigrants are defined as US-born with two US-born parents. Children born in Puerto Rico or in a foreign-country are defined as foreign-born. If data is missing on one parent, the immigrant generation is determined by the parent for who data is available.

Response format: 1 = Foreign-born 1st generation
2 = US born 2nd generation
3 = US born 3^{rd+} generation

Source Variables: USBORN, DPE5, DSE6

Note: This variable will be missing if data is not present for at least one biological parent. Children with missing parent information on one parent but whose second parent is foreign-born will be classified as second-generation. Children with missing information on one parent but whose second parent is US-born will be classified as third-generation.

Warning: This variable should not be used in analyses where health insurance, employment or college education are the primary outcomes. In these analyses, Puerto Ricans and others born in US territories should be treated as US born. Please use the IMGEN_PR variable instead.

5.4 IMGEN_PR: Immigrant Generation - Puerto Rican Revision (Child)
This is a three category immigrant generation variable. First-generation children are defined as foreign-born with foreign-born parents. Second-generation children are defined as US-born with at least one foreign-born parent. Third-generation immigrants are defined as US-born with two US-born parents. Children and parents born in Puerto Rico or another US territory are defined as US born. If data is missing on one parent, the immigrant generation is determined by the parent for who data is available.

Response format: 1 = Foreign-born 1st generation
2 = US born 2nd generation
3 = US born 3rd+ generation

Source Variables: USBORN_PR, DPE5, DSE6

Note: This variable will be missing if data is not present for at least one biological parent. Children with missing parent information on one parent but whose second parent is foreign-born will be classified as second-generation. Children with missing information on one parent but whose second parent is US-born will be classified as third-generation.

5.7 LANG_PREF: Language Preference - Child (Child)
This is a two category variable identifying which language was preferred to be used for the baseline examination. This was indicated on the clinic exam checklist. .

Response format: 1 = English
0 = Spanish

Source Variables: CKC0

5.5 IMCHILD: Children of Immigrants (Child)
This is a two category immigrant generation variable identifying children born to immigrant parents.

Response format: 1 = 1st or 2nd generation child of immigrants
0 = US born 3rd+ generation

Source Variables: IMGEN

Warning: This variable should not be used in analyses where health insurance, employment or college education are the primary outcomes. In these analyses, Puerto Ricans and others born in US territories should be treated as US born. Please use the IMCHILD_PR variable instead.

5.6 IMCHILD_PR: Children of Immigrants – Puerto Rican Revision (Child)
This is a two category immigrant generation variable identifying children born to immigrant parents. In this classification, Puerto Ricans and children born in other US territories are treated as US born.

Response format: 1 = 1st or 2nd generation child of immigrants
0 = US born 3rd+ generation

Source Variables: IMGEN_PR

5.8 IMMAGE_PR: Age at Migration (Child)
This is a categorical variable grouping children into those who migrated prior to school entry (ages 0-5), after school entry (ages 6-12), and during adolescence (ages 13+). In this classification, Puerto Ricans and children born in other US territories are treated as US born. Those migrating between the ages of 0-5 are often labeled the 1.75 generation. Those migrating between ages 6-12 are often labeled the 1.5 generation. Those migrating between ages 13-17 are labeled the 1.25 generation.

Response format: 1 = Migrated between ages 0-5
2 = Migrated between ages 6-12
3 = Migrated at age 13-17

Source Variables: PME1

Source Reference: Rubén G. Rumbaut. Ages, Life Stages, and Generational Cohorts: Decomposing the Immigrant First and Second Generations in the United States. International Migration Review, Vol. 38, No. 3, pp. 1160-1205, Fall 2004, Stable URL: http://www.jstor.org/stable/27645429

5.9 ACCSTRS: Acculturative Stress Index (Child)
This is a numeric variable. It is calculated as the average of the items in the scale and should range from of two category immigrant generation variable identifying children born to immigrant parents. If more than 2 of the 9 items are missing the score is missing.

Response format: 1-5

Source Variables: ASE items 1-9

Chronbach's Alpha (HCHS-Y): TBD

Source Reference: Andres G. Gil and William A. Vega Two Different Worlds: Acculturation Stress and Adaptation among Cuban and Nicaraguan Families. Journal of Social and Personal Relationships 1996 13: 435 DOI: 10.1177/0265407596133008. See also Gil, et al. Acculturation, familism, And alcohol use among Latino adolescent males: Longitudinal relations. Journal of Community Psychology, Vol. 28, No. 4, 443–458 (2000).

Note: Original paper does not utilize a score that sums across all items but divides the items into three categories – language conflicts, acculturation conflicts, perceived discrimination. Some analyses utilizing this scale appear to utilize sums. Other analyses appear to utilize averages. Different versions of the scale also include a 4-point and a 5-point Likert scale version. Some versions range from 0 to 4. No specific references indicate how to calculate the score with missing items.

5.10 ACCSTRS_L: Acculturative Stress Index: Language Conflict (Child)

This variable is calculated as the average of the two language conflict items in the acculturative stress scale. Higher values indicate greater language conflict. If either item is missing, the average is score is missing.

Response Format: Numeric ranging from 1-5

Source Variables: ASE items 1-2

Item Correlation (HCHS-Y): TBD

Source Reference: Andres G. Gil and William A. Vega Two Different Worlds: Acculturation Stress and Adaptation among Cuban and Nicaraguan Families. Journal of Social and Personal Relationships 1996 13: 435 DOI: 10.1177/0265407596133008. See also Gil, et al. Acculturation, familism, And alcohol use among Latino adolescent males: Longitudinal relations. Journal of Community Psychology, Vol. 28, No. 4, 443–458 (2000).

Note: Chronbach alpha's should not be calculated for an average of two items. A simple correlation coefficient should be calculated instead.

5.11 ACCSTRS_A: Acculturative Stress Index: Acculturation Conflicts (Child)

This variable is calculated as the average of the four conflict items in the acculturative stress scale. Higher values indicate greater acculturation conflict. If more than 1 of the 5 items is missing the score is missing.

Response Format: Numeric ranging from 1-5

Source Variables: ASE items 3-6

Chronbach's Alpha (HCHS-Y): TBD

Source Reference: Andres G. Gil and William A. Vega Two Different Worlds: Acculturation Stress and Adaptation among Cuban and Nicaraguan Families. Journal of Social and Personal Relationships 1996 13: 435 DOI: 10.1177/0265407596133008. See also Gil, et al. Acculturation, familism, And alcohol use among Latino adolescent males: Longitudinal relations. Journal of Community Psychology, Vol. 28, No. 4, 443–458 (2000).

5.12 ACCSTRS_D: Acculturative Stress Index: Discrimination (Child)
This variable is calculated as the average of the three discrimination items in the acculturative stress scale. Higher values indicate greater discrimination. If any one of the 3 items is missing the score is missing.

Response Format: Numeric ranging from 1-5

Source Variables: ASE items 7-9

Chronbach's Alpha (HCHS-Y): TBD

Source Reference: Andres G. Gil and William A. Vega Two Different Worlds: Acculturation Stress and Adaptation among Cuban and Nicaraguan Families. Journal of Social and Personal Relationships 1996 13: 435. DOI: 10.1177/0265407596133008. See also Gil, et al. Acculturation, familism, And alcohol use among Latino adolescent males: Longitudinal relations. Journal of Community Psychology, Vol. 28, No. 4, 443–458 (2000).

5.13 ARSMAII_MOS: Brief Acculturation Scale: Mexican-Orientation (Child)
This variable is calculated as the average of the six items in the Mexican-Orientation Scale (MOS). Higher values indicate a stronger Latino/Hispanic orientation. If 2 or more of the 6 items are missing the score is missing.

Response Format: Numeric ranging from 1-5

Source Variables: BAE items 1, 3, 6, 7, 8, 11

Chronbach's Alpha (HCHS-Y): TBD

Source Reference: Sheri Bauman. The Reliability and Validity of the Brief Acculturation Rating Scale for Mexican Americans-II for Children and Adolescents. Hispanic Journal of Behavioral Sciences 2005; 27; 426. DOI: 10.1177/0739986305281423

Warning: See the "Preliminary Psychometric Report on Acculturation Measures" for measure's properties and recommendations.

5.14 ARSMAII_AOS: Brief Acculturation Scale: American Orientation (Child)
This variable is calculated as the average of the six items in the American-Orientation Scale (AOS). Higher values indicate a stronger Latino/Hispanic orientation. If two or more of the 6 items is missing the score is missing.

Response Format: Numeric ranging from 1-5

Source Variables: BAE items 2,4,5,9,10,12

Chronbach's Alpha (HCHS-Y): TBD

Source Reference: Sheri Bauman. The Reliability and Validity of the Brief Acculturation Rating Scale for Mexican Americans-II for Children and Adolescents. Hispanic Journal of Behavioral Sciences 2005; 27; 426. DOI: 10.1177/0739986305281423

Warning: See the "Preliminary Psychometric Report on Acculturation Measures" for measure's properties and recommendations.

5.15 ARSMAII: Brief Acculturation Scale (Child)
This is a linear acculturation score derived by taking the mean of AOS minus the mean of MOS variable is calculated as the average of the six items in the American-Orientation Scale (AOS). Positive scores indicate an American orientation. Negative scores indicate a Hispanic/Latino orientation.
Response Format: Numeric ranging from -4 to 4

Source Variables: ARSMAII_AOS, ARSMAII_MOS

Source Reference: Sheri Bauman. The Reliability and Validity of the Brief Acculturation Rating Scale for Mexican Americans-II for Children and Adolescents. Hispanic Journal of Behavioral Sciences 2005; 27; 426. DOI: 10.1177/0739986305281423

Warning: See the "Preliminary Psychometric Report on Acculturation Measures" for measure's properties and recommendations.

5.16 ARSMAII_CATA: Brief Acculturation Scale Categories A (Child)
This is a categorical variable derived from the MOS and AOS scores. Cutpoints are based on standard deviations around the means of MOS and AOS scores. .

Response Format:
1 = Traditionalist (MOS $\geq$ 3.7 & AOS $\leq$ 3.24)
2 = Low Biculturals/Marginalized (MOS < 3.59 & AOS < 3.7)
3 = High Biculturals (MOS > 3.59 & AOS> 3.7)
4 = Assimilated (MOS$\leq$ 2.44 & AOS$\geq$ 4.11)
5 = Unclassified (those that don't fall into these other categories)

Source Variables: ARSMAII_AOS, ARSMAII_MOS

Source Reference: Sheri Bauman. The Reliability and Validity of the Brief Acculturation Rating Scale for Mexican Americans-II for Children and Adolescents. Hispanic Journal of Behavioral Sciences 2005; 27; 426. DOI: 10.1177/0739986305281423; See also Cuellar 1995 Acculturation Rating Scale for Mexican Americans-II. Cuellar's cutpoints differ somewhat from Bauman's.

Warning: See the "Preliminary Psychometric Report on Acculturation Measures" for measure's properties and recommendations.

5.16 ARSMAII_CATB: Brief Acculturation Scale Categories B (Child)
This is a categorical variable derived from the ARSMA II score.

Response Format:
1 = Mexican oriented (ARSMAII< -1.33)
2 = Mexican oriented bicultural (ARSMAII >= -1.33 & <= -.07)
3 = Anglo oriented bicultural (ARSMAII > -.07 & < 1.19)
4 = Strongly Anglo oriented (ARSMAII >=1.19 & < =2.45)
5= Very Assimilated (ARSMAII >2.45)

Source Variables: ARSMAII

Source Reference: Sheri Bauman. The Reliability and Validity of the Brief Acculturation Rating Scale for Mexican Americans-II for Children and Adolescents. Hispanic Journal of Behavioral Sciences 2005; 27; 426. DOI: 10.1177/0739986305281423; See also Cuellar 1995 Acculturation Rating Scale for Mexican Americans-II. Cuellar's cutpoints differ somewhat from Bauman's.

Warning: See the "Preliminary Psychometric Report on Acculturation Measures" for measure's properties and recommendations.

5.16 ARSMAII_CATC: Brief Acculturation Scale Categories C (Child)
This is a categorical variable derived from the ARSMA II score. It collapses ARSMAII_CATB into fewer categories.

Response Format:
1 = Mexican oriented (ARSMAII< -1.33)
2 = Bicultural (ARSMAII >= -1.33 & < 1.19)
4 = Anglo oriented (ARSMAII >=1.19)

Source Variables: ARSMAII_CATB

Source Reference: Sheri Bauman. The Reliability and Validity of the Brief Acculturation Rating Scale for Mexican Americans-II for Children and Adolescents. Hispanic Journal of Behavioral Sciences 2005; 27; 426. DOI: 10.1177/0739986305281423; See also Cuellar 1995 Acculturation Rating Scale for Mexican Americans-II. Cuellar's cutpoints differ somewhat from Bauman's.

Warning: See the "Preliminary Psychometric Report on Acculturation Measures" for measure's properties and recommendations.

5.17 MEIM_R: Ethnic Affirmation and Belonging (Child)
This is a numeric variable derived from 5 items from the MEIM-Revised Version that loaded most strongly on the ethnic affirmation and belonging subscale. The items are averaged together. If more than two items are missing, the score is set to missing. Higher numbers indicate a stronger sense of ethnic identity or, as discussed by Guarnaccia et al. (2007), enculturation.

Response Format: Values range from 1-5

Source Variables: EAE items 1-4 & 8

Chronbach's Alpha (HCHS-Y): TBD

Source Reference(s):
Roberts et al. The Structure of Ethnic Identity of Young Adolescents from Diverse Ethnocultural Groups. The Journal of Early Adolescence 1999 19: 301 DOI: 10.1177/0272431699019003001.

Guarnaccia et al. Assessing Diversity among Latinos: Results from the NLAAS. Hisp J Behav Sci. 2007; 29(4): 510–534. doi:10.1177/0739986307308110.

Notes: The standard MEIM is a 4-point scale ranging from 1-4. HCHS-youth adopted a 5-point scale for consistency with other scales in the study.

Warning: See the "Preliminary Psychometric Report on Acculturation Measures" for measure's properties and recommendations.

5.18 MMRI_C: Ethnic Centrality (Child)
This is a numeric variable derived from 4 items from the ethnic centrality subscale of the Multidimensional Model of Racial Identity (MMRI). Items in the scale are average together. If more than 1 item is missing, the score is set to missing. Higher numbers indicate that ethnic identity is more central to the child's identity.

Response Format: Values range from 1-5

Source Variables: EAE items 1,4,5, 6

Chronbach's Alpha (HCHS-Y): TBD

Source Reference:
Sellers et al. Multidimensional Model of Racial Identity: A Reconceptualization of African American Racial Identity Personality and Social Psychology Review 1998, Vol. 2, No. 1, 18-39.

Johnson et al. The Multidimensional Inventory of Black Identity: Its Use With Euro-American, Latino, and Native American Undergraduates. Measurement and Evaluation in Counseling and Development, Vol. 38, July 2005.

Kiang et al. Ethnic Identity and the Daily Psychological Well-Being of Adolescents From Mexican and Chinese Backgrounds. Child Development. Volume 77, Issue 5, pages 1338–1350, September/October 2006. DOI: 10.1111/j.1467-8624.2006.00938.x

Warning: See the "Preliminary Psychometric Report on Acculturation Measures" for measure's properties and recommendations.

5.19 MMRI_R: Ethnic Regard (Child)

This is a numeric variable derived from 4 items from the ethnic centrality subscale of the Multidimensional Model of Racial Identity (MMRI). Items in the scale are average together. If more than 1 item is missing, the score is set to missing. Higher values indicate that the child has a higher regard for their ethnic group.

Response Format: Values range from 1-5

Source Variables: EAE items 2,3,7,8

Chronbach's Alpha (HCHS-Y): TBD

Source Reference:

Sellers et al. Multidimensional Model of Racial Identity: A Reconceptualization of African American Racial Identity Personality and Social Psychology Review 1998, Vol. 2, No. 1, 18-39.

Johnson et al. The Multidimensional Inventory of Black Identity: Its Use With Euro-American, Latino, and Native American Undergraduates. Measurement and Evaluation in Counseling and Development, Vol. 38, July 2005.

Kiang et al. Ethnic Identity and the Daily Psychological Well-Being of Adolescents From Mexican and Chinese Backgrounds. Child Development. Volume 77, Issue 5, pages 1338–1350, September/October 2006. DOI: 10.1111/j.1467-8624.2006.00938.x

Warning: See the "Preliminary Psychometric Report on Acculturation Measures" for measure's properties and recommendations.

5.20 ETHID: Ethnic Identity (Child)

This is a numeric variable averaging all 8 items on ethnic identity to create a total score. If more than 3 items are missing, the score is set to missing. Higher values indicate that a stronger sense of ethnic identity or enculturation in the child.

Response Format: Values range from 1-5

Source Variables: EAE items 1-8

Chronbach's Alpha (HCHS-Y): TBD

Source Reference: See above.

Warning: See the "Preliminary Psychometric Report on Acculturation Measures" for measure's properties and recommendations.

5.21 AHISMA_A: US Orientation or Assimilation Subscale of AHISMA (Child)
This is a numeric variable calculated as the sum of responses to the 8 items which are equal to 1=US. Higher values indicate greater assimilation to the US.

Response Format: Values range from 0-8

Source Variables: UNE items 1-8

Chronbach's Alpha (HCHS-Y): is not appropriate as a measure of reliability for count variables. Instead, a KR-20 should be used. The 4 AHISMA Scales are perfectly collinear. Therefore, one scale must be omitted as a comparison group.

Source Reference: Unger et al. The AHISMA Acculturation Scale: A New Measure of Acculturation for Adolescents in a Multicultural Society. Journal of Early Adolescence, Vol. 22 No. 3, August 2002 225-251.

Warning: See the "Preliminary Psychometric Report on Acculturation Measures" for measure's properties and recommendations.

5.22 AHISMA_S: Separation Subscale of AHISMA (Child)
This is a numeric variable calculated as the sum of responses to the 8 items which are equal to 2=Family's Country (i.e. a country other than the US). This could be confusing to children whose are third-generation and whose parents come from the US. The assumption is that children will only choose 2 if their parents come from a country other than the US. Higher values indicate greater separation

Response Format: Values range from 0-8

Source Variables: UNE items 1-8

Chronbach's Alpha (HCHS-Y): is not appropriate as a measure of reliability for count variables. Instead, a KR-20 should be used. The 4 AHISMA Subscales are perfectly collinear (i.e. sum to 8). Therefore, one scale must be omitted as a comparison group.

Source Reference: Unger et al. The AHIMSA Acculturation Scale: A New Measure of Acculturation for Adolescents in a Multicultural Society. Journal of Early Adolescence, Vol. 22 No. 3, August 2002 225-251.

Warning: See the "Preliminary Psychometric Report on Acculturation Measures" for measure's properties and recommendations.

5.23 AHISMA_I: Integration Subscale of AHISMA (Child)
This is a numeric variable calculated as the sum of responses to the 8 items which are equal to 3=Both. Higher values indicate more integration.

Response Format: Values range from 0-8

Source Variables: UNE items 1-8

Chronbach's Alpha (HCHS-Y): is not appropriate as a measure of reliability for count variables. Instead, a KR-20 should be used. The 4 AHISMA Subscales are perfectly collinear (sum to 8). Therefore, one scale must be omitted as a comparison group.

Source Reference: Unger et al. The AHIMSA Acculturation Scale: A New Measure of Acculturation for Adolescents in a Multicultural Society. Journal of Early Adolescence, Vol. 22 No. 3, August 2002 225-251.

Warning: See the "Preliminary Psychometric Report on Acculturation Measures" for measure's properties and recommendations.

5.24 AHISMA_M: Marginalization Subscale of AHISMA (Child)
This is a numeric variable calculated as the sum of responses to the 8 items which are equal to 4=Neither. Higher values indicate more marginalization.

Response Format: Values range from 0-8

Source Variables: UNE items 1-8

Chronbach's Alpha (HCHS-Y): is not appropriate as a measure of reliability for count variables. Instead, a KR-20 should be used. The 4 AHISMA Subscales are perfectly collinear (sum to 8). Therefore, one scale must be omitted as a comparison group in any regression.

Source Reference: Unger et al. The AHIMSA Acculturation Scale: A New Measure of Acculturation for Adolescents in a Multicultural Society. Journal of Early Adolescence, Vol. 22 No. 3, August 2002 225-251.

Warning: See the "Preliminary Psychometric Report on Acculturation Measures" for measure's properties and recommendations.

5.25 AHISMA: AHISMA 4-Category Classification (Child)
This is a categorical variable calculated by classifying respondents into groups based on the AHISMA subscale score with the greatest sum. If there are ties, then assign the respondent to the highest of the tied categories. For example, if AHISMA_I = AHISMA_A then assign the respondent to category 2. If AHISMA_I = AHISMA_S then assign respondent to category 3.

Response Format: 1 = Integrated (Both)
2 = Assimilated (US)
3 = Separated (Foreign-Country)

4 = Marginalized (Neither)

Source Variables: AHISMA_A, AHISMA_S, AHISMA_I, AHISMA_M,

Source Reference: Unger et al. The AHIMSA Acculturation Scale: A New Measure of Acculturation for Adolescents in a Multicultural Society. Journal of Early Adolescence, Vol. 22 No. 3, August 2002 225-251.

Warning: See the "Preliminary Psychometric Report on Acculturation Measures" for measure's properties and recommendations.

5.26 AHISMA_EQ: AHISMA Equal Scores Indicator Variable (Child)
This is an indicator variable coded to indicate if the two highest-ranked AHISMA subscale scores were equal.

Response Format: 1 = Equal Scores
0 = No Equal Scores

Source Variables: AHISMA_A, AHISMA_S, AHISMA_I, AHISMA_M,

Chronbach's Alpha (HCHS-Y): Chronbach's Alpha is not appropriate as a measure of reliability for count variables. The 4 AHISMA Subscales are perfectly collinear (sum to 8). Therefore, one scale must be omitted as a comparison group.

Source Reference: Unger et al. The AHIMSA Acculturation Scale: A New Measure of Acculturation for Adolescents in a Multicultural Society. Journal of Early Adolescence, Vol. 22 No. 3, August 2002 225-251.

.
Warning: See the "Preliminary Psychometric Report on Acculturation Measures" for measure's properties and recommendations.

6. FAMILY FUNCTIONING and SOCIAL SUPPORT

6.1 PRFAMFTN: Poor Family Functioning (Child)

This scale is the 12-item General Functioning (GF) subscale of the McMaster Family Assessment Device (FAD). It provides a measure of the overall health of the family. Higher scores indicate poorer family functioning. The score is calculated as the average of items.[1] If 3 or more items are missing, the score is missing. Items 1, 3, 5, 7, 9, 11 must be reverse coded. In previous research, a score of 2.17 or greater (i.e. 10 percent of families) has been used to identify families with pathology.

Response Format: Values range from 1-4.

Source Variables: FFE items 1-12

Chronbach's Alpha (HCHS-Y): TBD

Source Reference: Byles et al. Ontario Family Health Study, General Functioning Subscale of the McMaster Family Assessment Device. Fam Proc 27:97-104, 1988

6.5 SOCSUPP: Social Support from Friends (Child)

This scale consists of 4 items from friendship subscale of the 12-item Multidimensional Scale of Perceived Social Support (MSPSS). A score is calculated as the average of the 4 items. If more than one item is missing, the score is set to missing. Higher values indicate more social support from friends.

Response Format: Values range from 1-4

Source Variables: SSE items 1-4

Chronbach's Alpha (HCHS-Y): TBD

Source Reference:
Zimet, G.D., Dahlem, N.W., Zimet, S.G. & Farley, G.K. (1988). The Multidimensional Scale of Perceived Social Support. Journal of Personality Assessment, 52, 30-41.

Zimet, G.D., Powell, S.S., Farley, G.K., Werkman, S. & Berkoff, K.A. (1990). Psychometric characteristics of the Multidimensional Scale of Perceived Social Support. Journal of Personality Assessment, 55, 610-17.

[1] Instructions indicate that the score should be calculated as the sum of items divided by 12. However, doing so could lead to a score of 0 when all items or missing. More generally, this calculation method leads to lower scores due to missing items. The instructions do not clearly indicate how to treat missing items.

Edwards, L. Measuring Perceived Social Support in Mexican American Youth: Psychometric Properties of the Multidimensional Scale of Perceived Social Support. Hispanic Journal of Behavioral Sciences 2004 26: 187. DOI: 10.1177/0739986304264374

Note: The original scale is on a 7-point likert scale.

7. PSYCHOLOGICAL WELL-BEING

7.1 CDE_SCORE: Child Depression Inventory (CDE) Score - Overall
This scale measures the severity of depression symptoms in children and is a 10-item abbreviation of the Child Depression Inventory developed by Kovacs (1985). The items are summed. The scores should be set to missing if more than 3 items are missing. .

Response Format: Values range from 0-20.

Source Variables: CDE items 1-10.

Chronbach's Alpha (HCHS-Y): TBD

Source Reference: Davanzo, et al. Spanish Translation and Reliability Testing of the Child Depression Inventory. Child Psychiatry and Human Development. Vol. 35(1), Fall 2004 pp 75-92.

Kovacs M: The Children's Depression, Inventory (CDI). Psychopharmacol Bull 21: 995–998, 1985.

Note: The translation used in the HCHS-Youth study differs slightly from the translation in the reference.

7.6 CDE_T_INT: Child Depression Inventory T-Score Internal Std
This numeric variable is a standardized conversion of the Child Depression Inventory score first to a z-score calculated using the SOL Youth study unweighted mean and standard deviation. Then that z-score was converted to a T score with a mean of 50 and a standard deviation of 10.

Response Format: Numeric variable ranging from 0 to 100.

Source Variables: CDE_SCORE

Source Reference: See above. Kovacs M: The Children's Depression, Inventory (CDI). Psychopharmacol Bull 21: 995–998, 1985.

7.7 CDE_R_INT: Child Depression Inventory Risk - Internal Std (3-Level)
Using the internal standardized t-scores, scores of 60-65 indicate a moderate risk for depression, while a score above 65 indicate clinically significant risk of depression.

Response Format: 1=Not at risk for depression
2=Moderate risk for depression (T > 60)
3=Severe risk for depression (T>65)

Source Variables: CDE_T_INT

Source Reference: See above. Kovacs M: The Children's Depression, Inventory (CDI). Psychopharmacol Bull 21: 995–998, 1985.

7.6 CDE_T_EXT: Child Depression Inventory T-Score External Std

This numeric variable is a standardized conversion of the Child Depression Inventory score first to a z-score calculated using external CDI standard mean and standard deviation by age group (8-12 and 13+) and gender. Then that z-score was converted to a T score with a mean of 50 and a standard deviation of 10.

Response Format: Numeric variable ranging from 0 to 100.

Source Variables: CDE_SCORE

Source Reference: See above. Kovacs M: The Children's Depression, Inventory (CDI). Psychopharmacol Bull 21: 995–998, 1985.

7.7 CDE_R_EXT: Child Depression Inventory Risk - External Std (3-Level)

Using the external standardized t-scores for child depression, scores of 60-65 indicate a moderate risk for depression, while a score above 65 indicate clinically significant risk of depression.

Response Format:
1=Not at risk for depression
2=Moderate risk for depression (T > 60)
3=Severe risk for depression (T>65)

Source Variables: CDE_T_INT

Source Reference: See above. Kovacs M: The Children's Depression, Inventory (CDI). Psychopharmacol Bull 21: 995–998, 1985.

7.5 MASC: Multi-Dimensional Anxiety Scale

This MASC-10 assesses the presence of anxiety disorders in youth and is widely used as a screening tool. Raw scores range from 0 to 30. The scores should be set to missing if more than 3 items are missing. Scores are converted to T-scores based on a mean of 50 and a standard deviation of 10. T-scores of greater than or equal to 56 indicate that a child has above average anxiety levels and is at risk for an anxiety-related disorder.

Response Format: Values range from 0-30.

Source Variables: MAE items 1-10

Chronbach's Alpha (HCHS-Y): TBD

Source Reference: See above. March JS, Parker JD, Sullivan K, Stallings P, Conners CK (1997) The Multidimensional Anxiety Scale for Children (MASC): Factor structure, reliability and validity. J Am Acad Child Adolesc Psychiatry. 36:554 –565.

7.6 MASC_T_INT: Multi-Dimensional Anxiety Scale T-Score Internal Std

This numeric variable is a standardized conversion of the MASC-10 to a T score first to a z-score calculated using the SOL Youth study unweighted mean and standard deviation. Then that z-score was converted to a T score with a mean of 50 and a standard deviation of 10.

Response Format: Numeric variable ranging from 0 to 100.

Source Variables: MASC

Source Reference: See above. March JS, Parker JD, Sullivan K, Stallings P, Conners CK (1997) The Multidimensional Anxiety Scale for Children (MASC): Factor structure, reliability and validity. J Am Acad Child Adolesc Psychiatry. 36:554 –565.

7.7 MASC_R_INT: Multi-Dimensional Anxiety Scale – Internal Std (Indicator >=56)

Using the internal standardized t-scores, a score at or above 56 indicate clinically significant risk of anxiety disorder..

Response Format: 0=Not at risk for anxiety disorder
1=At risk for anxiety disorder (MASC_T >= 56)

Source Variables: MASC_T

Source Reference: See above. March JS, Parker JD, Sullivan K, Stallings P, Conners CK (1997) The Multidimensional Anxiety Scale for Children (MASC): Factor structure, reliability and validity. J Am Acad Child Adolesc Psychiatry. 36:554 –565.

7.6 MASC_T_EXT: Multi-Dimensional Anxiety Scale

This numeric variable is a standardized conversion of the MASC-10 to a T score first to a z-score calculated using external MASC standard mean and standard deviation by age group (8-11, 12-15, and 16+) and gender. Then that z-score was converted to a T score with a mean of 50 and a standard deviation of 10.

Response Format: Numeric variable ranging from 0 to 100.

Source Variables: MASC

Source Reference: See above. March JS, Parker JD, Sullivan K, Stallings P, Conners CK (1997) The Multidimensional Anxiety Scale for Children (MASC): Factor structure, reliability and validity. J Am Acad Child Adolesc Psychiatry. 36:554 –565.

7.7 MASC_R_EXT: Multi-Dimensional Anxiety Scale Risk – External Std (Indicator>=56)
Using the external standardized t-scores, a score at or above 56 indicate clinically significant risk of anxiety disorder..

Response Format: 0=Not at risk for anxiety disorder
1=At risk for anxiety disorder (MASC_T >= 56)

Source Variables: MASC_T

Source Reference: See above. March JS, Parker JD, Sullivan K, Stallings P, Conners CK (1997) The Multidimensional Anxiety Scale for Children (MASC): Factor structure, reliability and validity. J Am Acad Child Adolesc Psychiatry. 36:554 –565.

7.8 LSE_SCORE: How I Feel Scale
This scale measures the social desirability of responses and is adapted from the Revised Children's Manifest Anxiety Scale developed by Reynolds and Paget. It is scored as the sum of responses to eight items.

Response Format: Values range from 0 to 8

Source Variables: LSE items 1-8

Chronbach's Alpha (HCHS-Y): TBD

Source Reference: Reynolds CR, Paget KO. National normative and reliability data for the Revised Children's Manifest Anxiety Scale. School Psyc Rev. 1983;12:324-336.

Note: Chronbach's Alpha is not appropriate for a count variable.

8. ANTHROPOMETRY

8.1 HEIGHT: Height corrected in cm

This is a numeric variable. Missing if all of the variables, ANT5a, ANT5b, and ANT5c were originally missing.

HEIGHT= average(ANT5a, ANT5b, ANT5c)

Source variables: ANT5a, ANT5b, ANT5c

8.2 BMI: Body Mass Index kg/m^2

This is a numeric variable. Missing if the variable, HEIGHT, is missing.

If HEIGHT is not missing then BMI = ANT6 / $(HEIGHT/100)^2$

Response Format: Values range from 10 to 80

Source variables: HEIGHT, ANT6.

8.3 BMIPCT: Body Mass Index Percentiles – CDC (Child)

Standardized SAS program obtained from CDC was used to calculate the percentiles and z-scores (standard deviations) for a child's sex and age for BMI, weight, and height based on the CDC growth charts (as of March 26, 2014).

Response Format: Values range from 0 to 100

Source variables: BMI

Source Reference: A SAS Program for the 2000 CDC Growth Charts (ages 0 to <20 y) from following web site (site date: 3/26/2014):
http://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm

8.4 BMIPCT95: Body Mass Index 95th Percentile – CDC (Child)

Standardized SAS program obtained from CDC was used to calculate the percentiles and z-scores (standard deviations) for a child's sex and age for BMI, weight, and height based on the CDC growth charts (as of March 26, 2014). This variable represents the 95th percentile for BMI.

Source variables: BMI

Source Reference: A SAS Program for the 2000 CDC Growth Charts (ages 0 to <20 y) from following web site (site date: 3/26/2014):
http://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm

8.5 BMIGRP_C5: Body Mass Index Percentile Groups (Child)

This is a numeric variable that uses the CDC standardized SAS program to calculate the percentiles and z-scores (standard deviations) for a child's sex and age for BMI, weight, and height based on the CDC growth charts (as of March 26, 2014).

"The use of the LMS parameters of the CDC growth charts has been shown to result in inaccurate estimates of the empirical percentiles at very high BMI values (e.g., the 99th percentile). Therefore, rather than using the BMI-for-age percentiles (and z-scores) to identify and track children who are extremely obese, it is recommended that these high BMI values be expressed as a percentage of the 95th percentile. A BMI value that is 20% greater than the 95th percentile (relative to the CDC reference population) is approximately equal to the 99th percentile of the reference population."

For example, a child with "a bmipct95 of 100 is at the 95th percentile of BMI-for-age. A value of 120 would indicate that the child's BMI is 20% greater than the 95th percentile." A final criterion of BMI greater than/equal to 35 imposed for severe obesity definition as well.

Response format:
1=underweight (BMI <5th percentile)
2=normal weight (BMI 5-84th percentile)
3=overweight (BMI 85-94th percentile)
4=obesity (BMI 95+ pctile and BMI < 35 and < 120% of 95%ile)
5=severe obesity (BMI 95+ pctile and BMI >=35 or >=120% of 95%ile)

Source variables: BMIPCT, BMIPCT95, AGE, GENDER

Source Reference: See above for A SAS Program for the 2000 CDC Growth Charts (ages 0 to <20 y) from following web site (site date: 3/26/2014):
http://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm

8.6 WAPCT: Weight-For-Age Percentile

This is a numeric variable that uses the standardized SAS program obtained from CDC was used to calculate the percentiles and z-scores (standard deviations) for a child's sex and age for BMI, weight, height, and head circumference based on the CDC growth charts."

Source variables: ANT6

Source Reference: See above for A SAS Program for the 2000 CDC Growth Charts (ages 0 to <20 y) from following web site (site date: 3/26/2014):
http://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm

8.7 WAIST: Waist Girth(Cm)

This is a numeric variable that is missing if all of the variables, ANT12a1, ANT12b1, ANT12c1 were originally missing.

WAIST= average(ANT12a1, ANT12b1, ANT12c1)

Source variables: ANT12a1, ANT12b1, ANT12c1

8.8 P90_WC: 90th %ile for Waist Circumference

This is a numeric variable that uses the 90th percentile for a child's sex and age for waist circumference based on the CDC growth charts (as of October 2012).

Response Format: Values range from 0 to 100

Source variables: AGE, GENDER.

Source reference: Fryar CD, Gu Q, Ogden CL. Anthropometric reference data for children and adults: United States, 2007–2010. National Center for Health Statistics. Vital Health Stat 11(252). 2012

8.9 CENTRAL_ADIPOSITY: Central Adiposity-CDC (2007-2010)

This is an indicator variable that uses the 90th percentile for a child's sex and age for waist circumference compared to their measured waist circumference to determine potential central (abdomen) adiposity.

WAIST >= P90_WC then CENTRAL_ADIPOSTY=1.
Otherwise if 0 <= WAIST <P90_WC then CENTRAL_ADIPOSTY=0.

Response Format: 1 = At risk for central (abdomen) adiposity (waist circ. >=90%ile)
0 = Not at risk for central (abdomen) adiposity

Source variables: WAIST, P90_WC.

Source reference: Fryar CD, Gu Q, Ogden CL. Anthropometric reference data for children and adults: United States, 2007–2010. National Center for Health Statistics. Vital Health Stat 11(252). 2012

9. BLOOD PRESSURE MEASURES

9.1 SBP_PCTILE: Systolic Blood Pressure Percentile - NHLBI

This is a numeric variable that calculates the percentiles from a child's height percentile z-score, sex and age for systolic blood pressure based on the NHLBI blood pressure level tables, which have been included in the following SAS code formulas (note: HAZ is height for age z-score obtained as a temporary variable from the CDC standardized SAS code used w/ anthropometric measurements such as BMI percentiles):

```
If gender='B' then do;

EXP_SBP=102.19768 + 1.82416*(AGE-10) + 0.12776*(AGE-10)**2 + 0.00249*(AGE-10)**3
+ -0.00135*(AGE-10)**4 + 2.73157*(HAZ) + -0.19618*(HAZ)**2 + -0.04659*(HAZ)**3 +
0.00947*(HAZ)**4;

SBP_ZSCORE=(SBP5-EXP_SBP)/10.7128;
end;

If gender='G' then do;
EXP_SBP=102.01027 + 1.94397*(AGE-10) +  0.00598*(AGE-10)**2 + -0.00789*(AGE-
10)**3 + -0.00059*(AGE-10)**4 + 2.03526*(HAZ) +  0.02534*(HAZ)**2 + -
0.01884*(HAZ)**3 +  0.00121*(HAZ)**4;

SBP_ZSCORE=(SBP5-EXP_SBP)/10.4855;
end;

SBP_PCTILE=PROBNORM(SBP_ZSCORE)*100;
```

Source variables: SBP5, AGE, GENDER, HAZ

Source reference: "The Fourth Report on the Diagnosis, Evaluation, and Treatment of High Blood Pressure in Children and Adolescents" Pediatrics. 2004 Aug;114 (2 Suppl 4th Report):555-76.

9.2 DBP_PNTILE: Diastolic Blood Pressure Percentile - NHLBI

This is a numeric variable that calculates the percentiles from a child's height percentile z-score, sex and age for diastolic blood pressure based on the NHLBI blood pressure level tables, which have been included in the following SAS code formulas (note: HAZ is height for age z-score obtained as a temporary variable from the CDC standardized SAS code used w/ anthropometric measurements such as BMI percentiles):

```
If gender='B' then do;
EXP_DBP=61.01217 + 0.68314*(AGE-10) + -0.09835*(AGE-10)**2 + 0.01711*(AGE-10)**3
+ 0.00045*(AGE-10)**4 + 1.46993*(HAZ) + -0.07849*(HAZ)**2 + -0.03144*(HAZ)**3 +
0.00967*(HAZ)**4;

DBP_ZSCORE=(SBP6-EXP_DBP)/11.6032;
```

end;

If gender='G' then do;
EXP_DBP=60.50510 + 1.01301*(AGE-10) + 0.01157*(AGE-10)**2 + 0.00424*(AGE-10)**3 + -0.00137*(AGE-10)**4 + 1.16641*(HAZ) + 0.12795*(HAZ)**2 + -0.03869*(HAZ)**3 + -0.00079*(HAZ)**4;

DBP_ZSCORE=(SBP6-EXP_DBP)/10.9573;
end;

DBP_PCTILE=PROBNORM(DBP_ZSCORE)*100;

Source variables: SBP6, AGE, GENDER, HAZ

Source reference: see above "The Fourth Report on the Diagnosis, Evaluation, and Treatment of High Blood Pressure in Children and Adolescents" Pediatrics. 2004 Aug;114 (2 Suppl 4th Report):555-76.

9.3 SBP_FLAG: SBP >=90 percentile

This is a 0/1 variable that determines whether/not the child's systolic BP percentile is greater than or equal to the 90th %ile.

Response Format: 0=No (< 90%ile)
1=Yes (90%ile or greater)

Source variables: SBP_PCTILE

Source reference: see above "The Fourth Report on the Diagnosis, Evaluation, and Treatment of High Blood Pressure in Children and Adolescents" Pediatrics. 2004 Aug;114 (2 Suppl 4th Report):555-76.

9.4 DBP_FLAG: DBP >=90 percentile

This is a 0/1 variable that determines whether/not the child's diastolic BP percentile is greater than or equal to the 90th %ile.

Response Format: 0=No (< 90%ile)
1=Yes (90%ile or greater)

Source variables: DBP_PCTILE

Source reference: see above "The Fourth Report on the Diagnosis, Evaluation, and Treatment of High Blood Pressure in Children and Adolescents" Pediatrics. 2004 Aug;114 (2 Suppl 4th Report):555-76.

9.5 HYPERTENSION_C3: Hypertension - NHLBI (Child)

This is a numeric variable that uses a child's systolic and diastolic blood pressure percentiles to determine a child's risk for hypertension based on the NHLBI blood pressure level tables. This variable does NOT include self-report doctor's diagnosis of hypertension or self-report antihypertensive medication use.

Response Format: 1=Normal (< 90%ile)
2=Prehypertensive (90-94%ile)
3=Hypertensive (>=95%ile)

Source variables: SBP_PCTILE, DBP_PCTILE

Source reference: see above "The Fourth Report on the Diagnosis, Evaluation, and Treatment of High Blood Pressure in Children and Adolescents" Pediatrics. 2004 Aug;114 (2 Suppl 4th Report):555-76.

10. LABORATORY MEASURES

10.1 DIABETES_GLUCOSE: Diabetes Based On Fasting Glucose >= 126

This is a 0/1 numeric variable that indicates diabetes diagnosis based on a fasting glucose value greater than or equal to 126 mg/dL.

If CHM14 >= 126 then DIABETES_GLUCOSE=1.
Else if 0 < CHM14 < 126 then DIABETES_GLUCOSE=0.

Response format: 0 = No
1 = Yes

Source variable(s): CHM14

Source reference:

10.2 DIABETES_TYPE2_SELF: Self-Report MD Diagnosis Of Type 2 Diabetes (Child)

This is a 0/1 numeric variable that indicates a self-report of diabetes. This variable is missing if MHE5 is missing.

if mhe5=1 and mhe5a=1 then DIABETES_TYPE2_SELF=1

else if mhe5=0 or (mhe5=1 & mhe5a^=1) then DIABETES_TYPE2_SELF=0

Response format: 0 = No
1 = Yes

Source variable(s): MHE5, MHE5a

Source reference:

10.3 HDL_FLAG: HDL <= 40 mg/dl

This is a 0/1 numeric variable that indicates a low HDL based on a value less than or equal to 40 mg/dL.

If 0 < CHM40 <= 40 then HDL_FLAG=1
Else if chm40 > 40 then HDL_FLAG=0

Response format: 0 = No
1 = Yes

Source variable(s): CHM40

Source reference: See above Cook S, Weitzman M, Auinger P, Nguyen M, Dietz WH. Prevalence of a metabolic syndrome phenotype in adolescents: Findings from the third National Health and Nutrition Examination Survey, 1988-1994.

10.4 IFG_IDF: Impaired Fasting Glucose - IDF

This is a 0/1 numeric variable that indicates impaired fasting glucose based on the IDF definition of metabolic syndrome in children and adolescents. This definition takes into account self-reported Type II diabetes, diabetic medication use, and a fasting glucose value greater than or equal to 100. This variable is missing if both fasting glucose and antidiabetic medication use indicators are missing.

If (CHM14 >= 100 OR DIABETES_TYPE2_SELF=1 OR MED_ANTIDIAB = 1) then IFG_IDF=1

Else IFG_IDF=0.

Response format: 0 = No
1 = Yes

Source variable(s): CHM14, DIABETES_TYPE2_SELF, MED_ANTIDIAB

Source reference: Alberti, K. G. M. M., Zimmet, P. and Shaw, J. (2006), Metabolic syndrome—a new world-wide definition. A Consensus Statement from the International Diabetes Federation. Diabetic Medicine, 23: 469–480.

Information packet can be obtained at: http://www.idf.org/metabolic-syndrome/children

10.5 LDL_FLAG: LDL >= 130 mg/dL

This is a 0/1 numeric variable that indicates a high LDL based on a value greater than or equal to 130 mg/dL.

If 0 < CHM53 >= 130 then LDL_FLAG=1.
Else if 0 < CHM53 < 130 then LDL_FLAG=0.

Response format: 0 = No
1 = Yes

Source variable(s): CHM53

Source reference: See above Cook S, Weitzman M, Auinger P, Nguyen M, Dietz WH. Prevalence of a metabolic syndrome phenotype in adolescents: Findings from the third National Health and Nutrition Examination Survey, 1988-1994.

10.6 TRIG_FLAG: Triglycerides >= 150 mg/dL
This is a 0/1 numeric variable that indicates high triglycerides based on a value greater than or equal to 150 mg/dL.

If CHM66 >= 150 then TRIG_FLAG=1;
Else if 0 < CHM66 < 150 then TRIG_FLAG=0.

Response format: 0 = No
1 = Yes

Source variable(s): CHM66

10.7 CHOL_FLAG: Cholesterol >= 200 mg/dL
This is a 0/1 numeric variable that indicates a high cholesterol based on a value greater than or equal to 200 mg/dL.

If CHM27 >= 200 then CHOL_FLAG=1.
Else if 0 < CHM27 < 200 then CHOL_FLAG=0.

Response format: 0 = No
1 = Yes

Source variable(s): CHM27

Source reference: See above Cook S, Weitzman M, Auinger P, Nguyen M, Dietz WH. Prevalence of a metabolic syndrome phenotype in adolescents: Findings from the third National Health and Nutrition Examination Survey, 1988-1994.

10.8 FPG_FLAG: Fasting Plasma Glucose >= 100 mg/dL
This is a 0/1 numeric variable that indicates a high fasting plasma glucose based on a value greater than or equal to 100 mg/dL.

If CHM14 >= 100 then FPG_FLAG=1
Else if 0 < CHM14 < 100 then FPG_FLAG=0.

Response format: 0 = No
1 = Yes

Source variable(s): CHM14

Source reference:

Source reference: See above Cook S, Weitzman M, Auinger P, Nguyen M, Dietz WH. Prevalence of a metabolic syndrome phenotype in adolescents: Findings from the third National Health and Nutrition Examination Survey, 1988-1994.

10.9 HBA1C_FLAG: HbA1C >= 5.7 mg/dl

This is a 0/1 numeric variable that indicates a high HbA1c based on a value greater than or equal to 5.7 mg/dL.

If LAB1 >= 5.7 then HBA1C_FLAG=1
Else if 0 < LAB1 < 5.7 then HBA1C_FLAG=0.

Response format: 0 = No
1 = Yes

Source variable(s): LAB1

10.10 HOMA_IR: HOMA index of Insulin Resistance

This is a numeric variable that calculates the HOMA index for insulin resistance by first calibrating the fasting plasma insulin (LAB11) from pmol/L to mU/mL in the formula (divide by 6). That value is then multiplied by fasting plasma glucose (LAB2), then dividing by the constant 405 (mass units constant).

For non-missing LAB2 and LAB11: LAB2*(LAB11/6)/405.

Source variable(s): LAB2, LAB11

10.11 INSULIN_RESIST: Insulin Resistance (HOMA_IR >= 2.5)

This is a 0/1 numeric variable that indicates a high insulin resistance based on a value greater than or equal to 2.5.

If HOMA_IR >= 2.5 then INSULIN_RESIST=1
Else if 0 < HOMA_IR < 2.5 then INSULIN_RESIST =0.

Response format: 0 = No
1 = Yes

Source variable(s): HOMA_IR

10.12 DYSLIPIDEMIA: Dyslipidemia (Child)

This is a 0/1 numeric variable that indicates a risk for dyslipidemia based on values of total cholesterol, LDL, HDL, and triglycerides. This variable is set to missing if all four cholesterol test values are missing.

If CHOL_FLAG=1 or LDL_FLAG=1 or TRIG_FLAG=1 or HDL_FLAG=1 then DYSLIPIDEMIA=1.
Else DYSLIPIDEMIA=0.

Response format: 0 = No
1 = Yes

Source variable(s): CHOL_FLAG, LDL_FLAG, TRIG_FLAG, HDL_FLAG

Source reference: See above Cook S, Weitzman M, Auinger P, Nguyen M, Dietz WH. Prevalence of a metabolic syndrome phenotype in adolescents: Findings from the third National Health and Nutrition Examination Survey, 1988-1994.

10.13 PREDIABETES: Diabetes using IFG > 100 and HbA1c > 5.7 (Child)
This is a 0/1 numeric variable that indicates a risk for prediabetes based on values of impaired fasting glucose and hbA1c. This variable is set to missing if both test values are missing.

If CHM14 > 100 OR CHM1 > 5.7 then set PREDIABETES=1;
Else if CHM14 = MISSING AND CHM1 = MISSING then set PREDIABETES = MISSING.
Otherwise, set PREDIABETES=0.

Response format: 0 = No
1 = Yes

Source variable(s): CMH1, CHM14

11. CLINICAL CHARACTERISTICS

11.1 MET_SYNDROME_ATP: Metabolic Syndrome - ATP-III (Child)

This is a 0/1 numeric variable with 1=metabolic syndrome present and 0=no metabolic syndrome.

The Adult Treatment Panel III (ATP III) of the National Cholesterol Education Program (NCEP) issued an evidence-based set of guidelines on cholesterol management in 2001 (*JAMA* 2001;285:2486-97). The following definition of metabolic syndrome is a modified child version based on the updated adult ATP III NCEP guidelines (*Circulation* 2004;109:433-438, *Circulation* 2004;110:227-239, http://www.americanheart.org/presenter.jhtml?identifier=4756).

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

BP_FLAG: SBP_ PCNTILE >= 90 or DBP_ PCNTILE >= 90 then BP_FLAG=1
TRIG_FLAG: TRIG_FLAG=1 (CHM66 >= 150)
HDL_FLAG: HDL_FLAG=1 (CHM40 <= 40)
FPG_FLAG: FPG_FLAG=1 (CHM14 >= 100)
DIABETES_TYPE2_SELF: DIABETES_TYPE2_SELF=1
CENTRAL_ADIPOSTY: CENTRAL_ADIPOSTY=1 (WAIST >= P90_WC: 90%ile)

First, each of the components is evaluated (0/1, or missing).
If all of the components are missing, then MET_SYNDROME_ATP should be missing.
Otherwise, MET_SYNDROME_ATP = 0 if sum of six components is 0, 1, 2
= 1 if sum of six components is 3, 4, 5, 6

Response format: 0 = No
1 = Yes

Source variables: AGE, GENDER, SBP_PCNTILE, DBP_PCNTILE, TRIG_FLAG, HDL_FLAG, FPG_FLAG, DIABETES_TYPE2_SELF, CENTRAL_ADIPOSTY

Source reference: see above "The Fourth Report on the Diagnosis, Evaluation, and Treatment of High Blood Pressure in Children and Adolescents" Pediatrics. 2004 Aug;114 (2 Suppl 4th Report):555-76.

Cook S, Weitzman M, Auinger P, Nguyen M, Dietz WH. Prevalence of a metabolic syndrome phenotype in adolescents: Findings from the third National Health and Nutrition Examination Survey, 1988-1994. *Arch Pediatr Adolesc Med.* 2003;157:821-827.

Chen W, Bao W, Begum S, Elkasabany A, Srinivasan SR, Berenson GS. Age-related patterns of the clustering of cardiovascular risk variables of syndrome X from childhood to young adulthood in a population made up of black and white subjects: The Bogalusa Heart Study. *Diabetes.* 2000;49:1042-1048.

Duncan GE, Li SM, Zhou XH. Prevalence and trends of a metabolic syndrome phenotype among US adolescents, 1999-2000. *Diabetes Care.* 2004;27:2438-2443.

Weiss R, Dziura J, Burgert TS, Tamborlane WV, Taksali SE, Yeckel CW, Allen K, Lopes M, Savoye M, Morrison J, Sherwin RS, Caprio S. Obesity and the metabolic syndrome in children and adolescents. *N Engl J Med.* 2004:350:2362-2374.

Julia Steinberger et al. Progress and Challenges in Metabolic Syndrome in Children and Adolescents: A Scientific Statement from the American Heart Association.. *Circulation.* 2009;119:628-647

11.2 MET_SYNDROME_IDF: Metabolic Syndrome - IDF (Child)

This is a 0/1 numeric variable with 1=metabolic syndrome present and 0=no metabolic syndrome.

The International Diabetes Federation (IDF) provides the following definition of metabolic syndrome in children and adolescents. The IDF definition emphasizes central obesity because it is "independently" associated with each of the other metabolic syndrome components, as well as insulin resistance. From the IDF Consensus Definition document:

"The new IDF definition is divided according to age-groups because of developmental challenges presented by age-related differences in children and adolescents: age 6 years to younger than 10 years; age 10 years to younger than 16 years; and 16 years or older. Children who are younger than 6 years were excluded from the definition because of insufficient data for this age-group. In all three age groups, abdominal obesity is the "sine qua non".

IDF suggests that the metabolic syndrome should not be diagnosed in children younger than 10 years, but that a strong message for weight reduction should be delivered for those with abdominal obesity.

For children age 10 years or older, metabolic syndrome can be diagnosed with abdominal obesity and the presence of two or more other clinical features (i.e. elevated triglycerides, low HDL-cholesterol, high blood pressure, increased plasma glucose). In the absence of contemporary definitive data, the criteria adhere to the absolute values in the IDF adult definition, except that waist circumference percentiles are recommended and one (rather than a sex-specific) cut-off is used for HDL- cholesterol levels. For children older than 16 years, the IDF adult criteria can be used."

The IDF diagnosis of metabolic syndrome for children is made:

a) CENTRAL_ADIPOSTY: CENTRAL_ADIPOSTY=1 (WAIST >= P90_WC: 90%ile) or WAIST ≥ 94 (boys), WAIST≥ 80 cm (girls)

b) Plus any 2 of the following:

   - BP_FLAG_2: SBP5 >= 130 (systolic) or SBP6 >= 85 (diastolic) then BP_FLAG_2=1

- TRIG_FLAG: TRIG_FLAG=1 (CHM66 >= 150)
- HDL_FLAG: HDL_FLAG=1 (CHM40 <= 40)
- FPG_FLAG: FPG_FLAG=1 (CHM14 >= 100)
- DIABETES_TYPE2_SELF: DIABETES_TYPE2_SELF=1

The IDF diagnosis of metabolic syndrome for adults (Youth age >=16) is made:

a) ABDOMINAL_OBESITY_IDF: WAIST ≥ 94 (boys), WAIST≥ 80 cm (girls)

b) Plus any 2 of the following:

- HIGH_TRIG_PLUSMEDS: CHM66≥150 mg/dL or treatment for this lipid abnormality (MED_ FIBARES_NICOACID =1)
- LOW_HDL_PLUSMEDS: CHM40< 40 mg/dL (boys), CHM40< 50 mg/dL (girls) or specified treatment for this lipid abnormality (MED_ FIBARES_NICOACID =1)
- ELEVATED_BP_PLUSMEDS: SBP5 ≥ 130 or SBP6 ≥ 85 mm Hg or use of anti-hypertension medications (MED_ANTIHYPERT =1) (note: requires non-missing SPB, DBP, or medication use).
- IFG_IDF: CHM14 ≥ 100 mg/dl, or previous diagnosis of diabetes (DIABETES_TYPE2_SELF =1), or use of anti-diabetic medications (MED_ANTIDIAB=1).
-

First, each of the components is evaluated (0/1, or missing).

If abdominal obesity is missing or 2 or more or the other components are missing components, then METS_IDF should be missing.

Otherwise, METS_IDF = 0 if abdominal obesity is 0 or abdominal obesity is 1 and sum of four remaining components is 0 or 1.
= 1 if abdominal obesity is 1 and the sum of the four remaining components is 2, 3, or 4.

Response format: 0 = No
1 = Yes

Source variables: AGE, GENDER, SBP_PCNTILE, DBP_PCNTILE, SBP5, SBP6, TRIG_FLAG, HDL_FLAG, FPG_FLAG, DIABETES_TYPE2_SELF, CENTRAL_ADIPOSTY, ABDOMINAL_OBESITY_IDF, HIGH_TRIG_PLUSMEDS, LOW_HDL_PLUSMEDS, MED_ FIBARES_NICOACID, ELEVATED_BP_PLUSMEDS, MED_ANTIHYPERT, IFG_IDF, MED_ANTIDIAB

Source reference: see above "The Fourth Report on the Diagnosis, Evaluation, and Treatment of High Blood Pressure in Children and Adolescents" Pediatrics. 2004 Aug;114 (2 Suppl 4th Report):555-76.

Alberti, K. G. M. M., Zimmet, P. and Shaw, J. (2006), Metabolic syndrome—a new world-wide definition. A Consensus Statement from the International Diabetes Federation. Diabetic Medicine, 23: 469–480.

Information packet can be obtained at: http://www.idf.org/metabolic-syndrome/children

11.3 CVD_RISK_COUNT: CVD Risk Factors Count

This is a numeric variable that counts the total number of CVD risk factors a child participant has based on the following conditions (using SUM function to allow for missing conditions):

OBESITY_BIN
SBP_FLAG or DBP_FLAG
FPG_FLAG
HBA1C_FLAG
CHOL_FLAG
LDL_FLAG
TRIG_FLAG
HDL_FLAG
INSULIN_RESIST

Source variable(s): OBESITY_BIN, SBP_FLAG, DBP_FLAG, FPG_FLAG, HBA1C_FLAG, CHOL_FLAG, LDL_FLAG, TRIG_FLAG, HDL_FLAG, INSULIN_RESIST

11.4 CVD_RISK_C4: CVD Risk Factors Count (0, 1, 2, 3+)

This is a 4-level numeric variable that indicates groups based on the total number of CVD risk factors a child participant, as follows:

If CVD_RISK_COUNT >=3 then set CVD_RISK_C4 = 3
Else CVD_RISK_C4 = CVD_RISK_COUNT

Response format:
0=0 risk factors
1=1 risk factor
2=2 risk factors
3=3 or more risk factors

Source variable(s): CVD_RISK_COUNT

11.5 CVD_RISK_C4: CVD Risk Factors Count (0, 1, 2, 3+)

This is a 4-level numeric variable that indicates groups based on the total number of CVD risk factors a child participant, as follows:

If CVD_RISK_COUNT >=3 then set CVD_RISK_C4 = 3

Else CVD_RISK_C4 = CVD_RISK_COUNT

Response format: 0=0 risk factors
1=1 risk factor
2=2 risk factors
3=3 or more risk factors

Source variable(s): CVD_RISK_COUNT

11.6 CVD_RISK_C3: CVD Risk Factors Count (0, 1-2, 3+)

This is a 3-level numeric variable that indicates groups based on the total number of CVD risk factors a child participant, as follows:

If CVD_RISK_COUNT >=3 then set CVD_RISK_C3 = 2
Else if CVD_RISK_COUNT in (1,2) then CVD_RISK_C3=1
else CVD_RISK_C3 = CVD_RISK_COUNT
Response format: 0=0 risk factors
1=1-2 risk factors
2=3 or more risk factors

Source variable(s): CVD_RISK_COUNT

11.7 CVD_RISK_BIN: CVD Risk Factors Count (0-2, 3+)

This is a 0/1 numeric variable that indicates groups based on the total number of CVD risk factors a child participant, as follows:

If CVD_RISK_COUNT >=3 then set CVD_RISK_BIN = 1
Else if CVD_RISK_COUNT in (0,1,2) then CVD_RISK_BIN=0

Response format: 0=0-2 risk factors
1=3 or more risk factors

Source variable(s): CVD_RISK_COUNT

11.8 PDE_TANNER_MAP: PDS Scoring Shirtcliff Version (Child)

This is a numeric variable that determines the stages of pubertal development using self-reported scores of the Pubertal Development Scale (PDE) mapped to the Tanner Staging groups. This mapping algorithm was created by Dr. Elizabeth Shirtcliff and obtained through correspondence with the SOL Youth study by request. Additional correspondence (6/16/2014) determined that in cases when adrenal or gonadal scores were in between groups (ex. 2.25, 3.5) that is was suggested that the score be increased to the next integer (ex. 2.25 to 3).

Three variables were created for the PDE to tanner Stage mapping, including individual PDE/Tanner stages for gonadal and adrenal separately, as well as a composite score combining both:

PDE_TANNER_MAP_A: PDS Adrenal

PDE_TANNER_MAP_G: PDS Gonadal

PDE_TANNER_MAP: PDS Scoring Shirtcliff Version

Source coding obtained was in SPSS format and the CSCC translated this into SAS coding, which is below (SPSS coding would be made available upon request):

```
     *-- Recode PDE1_-PDE4_,PDE6_,PDE7_;
     if gender='G' then do;
       if pde1 in(1,2,3,.) then PDE1_=PDE1;
       else if pde1=4 then PDE1_=PDE1+1;
     end;
     else if gender='B' then do;
       if pde1 in(1,.) then PDE1_=PDE1;
       else if pde1 in(2,3,4) then PDE1_=PDE1+1;
     end;
     label PDE1_='Growth in height recoded (PDE1 converted to 5 point
scale - dependent on gender)';

     if pde2 in(1,2,.) then PDE2_=PDE2;
     else if pde2 in(3,4) then PDE2_=PDE2+1;
     label PDE2_='Growth of body hair recoded (PDE2 converted to 5
point scale)';

     if gender='G' then do;
       if pde3 in(1,2,.) then PDE3_=PDE3;
       else if pde3 in(3,4) then PDE3_=PDE3+1;
     end;
     else if gender='B' then PDE3_=PDE3;
     label PDE3_='Notice skin changes recoded (PDE3 converted to 5
point scale - dependent on gender)';

     if gender='B' then do;
       if pde4 in(1,2,3,.) then PDE4_=PDE4;
       else if pde4 in(4) then PDE4_=PDE4+1;
     end;
     else if gender='G' then PDE4_=.;
     label PDE4_='Deepening of voice recoded (PDE4)';

     if gender='G' then do;
       if pde6 in(1,.) then PDE6_=PDE6;
       else if pde6 in(2,3,4) then PDE6_=PDE6+1;
     end;
     else if gender='B' then PDE6_=.;
```

```
     label PDE6_='Breasts begun to grow recoded (PDE6 converted to 5
point scale)';

     if gender='G' then do;
       if pde7 in(.) then PDE7_=PDE7;
       else if pde7 in(0) then PDE7_=1;
       else if pde7=1 then PDE7_=5;
     end;
     else if gender='B' then PDE7_=.;
     label PDE7_='Female begun to menstruate recoded (PDE7 converted
to 5 point scale)';

     *-- Define ADRENF;
     if gender='G' & pde2_>.Z & pde3_>.Z
       then ADRENF=mean(pde2_,pde3_);
     else ADRENF=.;
     label ADRENF='Female Adrenal Score(average of recoded PDE2 and
PDE3)';

     *-- Define ADRENF2;
     if gender='G' then do;
       if adrenf in(1,.) then ADRENF2=adrenf;
       else if pde2=1 & adrenf=1.5 then ADRENF2=1;
       else if pde2=2 & adrenf=1.5 then ADRENF2=2;
       else if 2<=adrenf<=5 then ADRENF2=round(adrenf);
       else if adrenf=5.5 then ADRENF2=5;
     end;
     else if gender='B' then ADRENF2=.;
     label ADRENF2='Adrenal score for girls based on skin changes and
pubic hair';

     *-- Define GONADF;
     if gender='G' & pde1_>.Z & pde6_>.Z
       then GONADF=mean(pde1_,pde6_);
     else GONADF=.;
     label GONADF='Female Gonadal Score(Mean of recoded PDE1 and
PDE6)';

     *-- Define GONADF2;
     if gender='G' then do;
       if pde7_=1 then do;
         if gonadf=. then GONADF2=.;
         else if 1<=gonadf<4 then GONADF2=floor(gonadf);
         else if gonadf=4 then GONADF2=3;
         else if gonadf in(4.5,5) then GONADF2=4;
       end;
       else if pde7_=5 then do;
         if gonadf=. then GONAF2=.;
         else if gonadf=1 then GONADF2=2;
         else if gonadf=1.5 then GONADF2=3;
         else if 2<=gonadf<=3 then GONADF2=4;
         else if 3<gonadf<=5 then GONADF2=5;
```

```
      end;
      else if pde7_<=.Z then GONADF2=gonadf;
    end;
    else if gender='B' then GONADF2=.;
    label GONADF2='Gonadal score for girls based on growth spurt,
breast development and menarche';

    *-- Define PDSS_F;
    if gender='G' & gonadf2>.Z & adrenf2>.Z then do;
       PDSS_F=mean(gonadf2,adrenf2);
       if PDSS_F in(1,2,3,4,5) then PDSS_F=PDSS_F;
       else PDSS_F=round((mean(gonadf2,adrenf2)+1));
       if PDSS_F>5 then PDSS_F=5;
    end;
    else PDSS_F=.;
    label PDSS_F='Female-PDS scoring shirtcliff version';

    *-- Define ADRENM;
    if gender='B' & pde2_>.Z & pde3_>.Z
      then ADRENM=mean(pde2_,pde3_);
    else ADRENM=.;
    label ADRENM='Male Adrenal Score (average of recoded PDE2 and
PDE3)';

    *-- Define ADRENM2;
    if gender='B' then do;
      if adrenm in(1,2,3,4,5) then ADRENM2=adrenm;
      else if (adrenm in(1.5) & pde3=1) | (adrenm=2.5 & pde3_<4) then
ADRENM2=floor(adrenm);
      else if (adrenm in(1.5) & pde3=2) | (adrenm=2.5 & pde3_=4) then
ADRENM2=ceil(adrenm);
      else if adrenm in(3.5,4.5) then ADRENM2=ceil(adrenm);
      else if adrenm=5.5 then adrenm2=5;
    end;
    label ADRENM2='Adrenal score for boys based on skin changes and
pubic hair';

    *-- Define GONADM;
    if gender='B' & pde1_>.Z & pde4_>.Z
       then GONADM=mean(pde1_,pde4_);
    else gonadm=.;
    label GONADM='Male Gonadal Score(Mean of recoded PDE1 and PDE4)';

    *-- Define GONADM2;
    if gender='B' then do;
      if pde5<=.Z & gonadm>.Z then GONADM2=gonadm;
      else if pde5=1 & gonadm in(1,4) then GONADM2=gonadm;
      else if pde5>1 & gonadm=1 then GONADM2=2;
      else if pde5=1 & gonadm in(1.5,2.5) then GONADM2=floor(gonadm);
      else if pde5>1 & gonadm in(1.5,2.5) then GONADM2=ceil(gonadm);
      else if gonadm=2 & pde5=1 & pde4=1 then GONADM2=1;
      else if gonadm=2 & pde5=1 & pde4>1 then GONADM2=2;
```

```
        else if gonadm=2 & pde5>1 then GONADM2=3;
        else if gonadm=3 then GONADM2=3;
        else if gonadm=3.5 & pde5 in(1,2) then GONADM2=4;
        else if gonadm=3.5 & pde5>2 then GONADM2=5;
        else if gonadm=4 & pde5=2 then GONADM2=4;
        else if gonadm=4 & pde5>2 then GONADM2=5;
        else if gonadm>4 then GONADM2=5;
        else if gonadm<=.Z then GONADM2=.;
      end;
      else if gender='G' then GONADM2=.;
      label GONADM2='Gonadal score for boys based on growth spurt,
deepening if voice and facial hair';

      *-- Define PDSS_M;
      if gender='B' & gonadm2>.Z & adrenm2>.Z then do;
         PDSS_M=mean(gonadm2,adrenm2);
         if PDSS_M in(1,2,3,4,5) then PDSS_M=PDSS_M;
         else PDSS_M=round((mean(gonadm2,adrenm2)+1));
         if PDSS_M>5 then PDSS_M=5;
      end;
      else PDSS_M=.;
      label PDSS_M='Male-PDS scoring shirtcliff version';

      *-- Define PDE_TANNER_MAP;
      if gender='B' then PDE_TANNER_MAP=PDSS_M;
      else if gender='G' then PDE_TANNER_MAP=PDSS_F;
      label PDE_TANNER_MAP='PDS scoring shirtcliff version';

      *-- Define PDE_TANNER_MAP_A;
      if gender='B' then PDE_TANNER_MAP_A=adrenm2;
      else if gender='G' then PDE_TANNER_MAP_A=adrenf2;
      label PDE_TANNER_MAP_A='PDS adrenal';

      *-- Define PDE_TANNER_MAP_G;
      if gender='B' then PDE_TANNER_MAP_G=ceil(gonadm2);
      else if gender='G' then PDE_TANNER_MAP_G=ceil(gonadf2);
      label PDE_TANNER_MAP_G='PDS gonadal';
```

Response format:
1= Stage I
2= Stage II
3= Stage III
4= Stage IV
5= Stage V

Source variables: Age, GENDER, PDE1-PDE7

Source Reference: Shirtcliff EA, Dahl RE, Pollak SD: Pubertal development: correspondence between hormonal and physical development. Child Dev 2009;80:327-337.

11.9 VO2_MAX: Estimated Maximum VO2 – step test (ml/kg/min)

This is a numeric variable that calculates the estimated maximum VO2 (aerobic fitness) based on the results from the fitness step test, as follows:

VO2_MAX = 105.3959 – (1.643756 x FST5)

Source variable(s): FST5

Source Reference: McMurray RG, Bangdiwala SI, Harrell JS, Amorim LD. Adolescents with metabolic syndrome have a history of low aerobic fitness and physical activity levels. Dyn Med. 2008;7:5. (in eng).

12. MEDICATIONS – SCANNED/TRANSCRIBED ONLY

All parent(s) were asked to bring in all of the child(ren)'s prescribed or over-the-counter medications taken in the past 4 weeks to the baseline examination for inventory. MUE1 and MUE2 are used to define individuals who did/did not bring in any medications as well as those who have/have not taken any medications in the past 4 weeks in order to determine absence of medication use. Medications are then scanned or transcribed into medication coding fields MUE5A-MUE5E to MUE14A-MUE14E (up to 10 possible medications). For the derived indicator variables below, participants were set to missing unless the reported MUE2=1 (took no medication) or reported at least 1 medication in the fields above.

The medication information is them compared to the HCHS master drug file which draws data from several sources:

1. Medi-Span's Master Drug Data Base (MDDB ®, version 2, 2003)
2. Lexi-Comp ® Online value-added version of Index Nominum Drug Data Base (version 2009)
3. Ovid's Martindale Complete Drug Reference (version 2009)

Mexican (MX) and Dominican (DO) medications from (2-3) above were extracted and then assigned therapeutic classification codes associated with their generic ingredients in (1). Assignments were automatic when data retrieval from (2-3) was fully electronic and generic ingredients in the source files matched exactly. Assignments were manual when data retrieval from (2-3) was semi-manual or generic ingredients in the source files did not match exactly and required review by a health professional.

The HCHS master drug file "contains basic drug names, National Drug Codes (NDCs), and assorted descriptive information used to uniquely identify drug products. From this data, generic product identifiers (GPI) are obtained which groups drug products by a hierarchical therapeutic classification scheme. The 14-character GPI consists of a hierarchy of seven subsets, each providing increasingly more specific information about drug products. These subsets are structured and identified below" (from MDDB ®, version 2, 2003):

GPI Subset	Record Type	Size	Representation	Example
12-xx-xx-xx-xx-xx-xx	1	2	Drug Group	*MISC. ENDOCRINE*
12-34-xx-xx-xx-xx-xx	2	4	Drug Class	*Posterior Pituitary**
12-34-56-xx-xx-xx-xx	3	6	Drug Subclass	*Vasopressin***
12-34-56-78-xx-xx-xx		8	Drug Name	Desmopressin
12-34-56-78-90-xx-xx	4	10	Drug Name	Acetate
12-34-56-78-90-12-xx		12	Dosage Form	Tablet
12-34-56-78-90-12-34	5	14	Strength	0.1MG

*NOTE: The below medications were indicated for use within the last 4 weeks from baseline examination.

12.1 Med_Antianginal (Antianginals)

This is a 0/1 numeric variable.

Medication Codes:
The “32xxxx” group of medications: *ANTIANGINAL AGENTS*

12.2 Med_Antiasthmatics (Antiasthmatic or bronchodilator agents)

This is a 0/1 numeric variable.

Medication Codes:
The “44xxxx” group of medications: *ANTIASTHMATIC AND BRONCHODILATOR AGENTS*

12.3 Med_Antidiab (Antidiabetics)

This is a 0/1 numeric variable.

Medication Codes:
Any “27xxxx” group of medications: *ANTIDIABETICS*

12.4 Med_Antihypert (Antihypertensives)

This is a 0/1 numeric variable.

Medication Codes:

Any “36xxxx” group of medications: *ANTIHYPERTENSIVES*

12.5 Med_Aspirin (Aspirin-containing analgesics)

This is a 0/1 numeric variable.

Medication Codes:
Any “6410xx” class of medications: *Salicylates**
Any “6420xx” class of medications: *Analgesics Other**
Any “6499xx” class of medications: *Analgesic Combinations**
Any “6599xx” class of medications: *Narcotic Combinations**

12.6 Med_BB (Beta Blockers)

This is a 0/1 numeric variable.

Medication Codes:
Any “33xxxx” group of medications: *BETA BLOCKERS*
Any “369920” subclass of medications: *Beta Blocker & Diuretic Combinations***
Any “369925” subclass of medications: *Beta Blocker & Calcium Channel Blocker Combinations***

12.7 Med_Chemo (Chemotherapy)

This is a 0/1 numeric variable.

Medication Codes:

Any “21xxxx” group of medications: *ANTINEOPLASTICS*

12.8 Med_Fibares_Nicoacid (Fibric/Nicotinic Acids (trt of TG and HDL))

This is a 0/1 numeric variable.

Medication Codes:
Any “392000” subclass of medications: *Fibric Acid Derivatives**
Any “394500” subclass of medications: *Nicotinic Acid Derivatives**
Any “399920” subclass of medications: *Fibric Acid Derivative Combinations***

12.9 Med_Metformin (Metformins)

This is a 0/1 numeric variable.

Medication Codes:
Any “2725xx” class of medications: *Biguanides**
Any “2799xx” class of medications: *Antidiabetic Combinations**

12.10 Med_NSAID (NSAIDs)

This is a 0/1 numeric variable.

Medication Codes:
Any “6610xx” class of medications: *Nonsteroidal Anti-inflammatory Agents (NSAIDs)**

12.11 Med_OI_Steroid (Oral/inhalable Glucocorticosteroids)

This is a 0/1 numeric variable. Does not include not including mineralocorticoids, androgens, estrogens, and progestogens.

Medication Codes:
The “221000” subclass of medications: *Glucocorticosteroids**
The “221099” subclass of medications: *Steroid Combinations***
The “444000” subclass of medications: *Steroid Inhalants**
The “449947” subclass of medications: *Sympathomimetic-Steroid w/ Anti-infective***
The “449980” subclass of medications: *Steroid-Antihistamines***
The “449982” subclass of medications: *Steroid-Antihistamine-Expectorants***
The “449984” subclass of medications: *Steroid-Sympathomimetic-Antihistamine-&/or Expectorant***
The “449988” subclass of medications: *Xanthine-Steroids***

12.12 Med_Statin (Statins)

This is a 0/1 numeric variable.

Medication Codes:
Any “3940xx” class of medications: *HMG CoA Reductase Inhibitors**

13.Physical Activity - Self-Report (derived from PAE q'aire)

The SOL Youth self-report physical activity form (PAE) includes a 68-item questionnaire that uses activity-based and time-based approaches for obtaining self-reported moderate-to-vigorous physical activity from adolescents and children. In addition asked 4 questions on how much time per day (in hours and minutes) children spent in certain sedentary behaviors. These were converted to create a number of times/month. See SOL Youth Child Derive Variable Dictionary for more detailed descriptions of the self-report questionnaire (PAE) derived variables.

The SOL Youth physical activity questionnaire (PAE) asked children to estimate how often they participated in each of 68 activities in the past month. The response options included never, 1-2 times/month, 1-2 times/week, 3-4 times/week, 5-6 times/week, and daily. These were converted to create a number of times/month. The 68 activities were grouped by mode based on the compendium of physical activities (Ridley K, Int J Behav Nutr Phys Act. 2008) into the following 5 categories: transport, school, leisure non-sport, sports / exercise, and household. Items assigned to each mode of activity were summed to derive number of times/month spent in each mode. In summary, there are 5 variables with times/month spent in each mode.

Each activity reported was assigned a metabolic equivalent (MET) value from the MET table used in a prior study (McMurray et al, Med Sci Sports Exerc. 2004) and from the compendium of physical activities 2. Activities can be grouped into 4 broad categories:
1.0-1.5 METS sedentary
1.6-2.9 METS light
3.0-5.9 METS moderate
>=6.0 METS vigorous

Activities listed as "Other" by a participant were reviewed to determine activity type and intensity.

The PAE questionnaire also asked how much time per day (in hours and minutes) children spent in the following 4 sedentary behaviors: television or video watching, computer or internet, video/computer games (non-active), and talking on the phone or text messaging. These activities were summed to create total time (in hours and minutes) spent in 4 sedentary behaviors: television or video watching, computer or internet, video/computer games (non-active), and talking on the phone or text messaging.

Variable	Label	Definition

mpae1 to mpae68	Times/month for pae1-pae68	Here, X=1 to 68. So repeat the following steps for pae1-pae68.
		if paeX=1 then mpaeX=0;
		if paeX=2 then mpaeX=1.5;
		if paeX=3 then mpaeX=6;
		if paeX=4 then mpaeX=14;
		if paeX=5 then mpaeX=22;
		if paeX=6 then mpaeX=30;

Variable	Label	Definition
PAE65a_sports PAE65a_household PAE65a_leisure	Recoding PAE65a into activity types	PAE65a is a blank, which asks for other activities. In this step, we try to re-categorize PAE65a into 5 activity types (transportation, school, leisure non sport, sports/exercise and household). See Table 1 for decisions of variables below and table 2 for hard-coding of variables based on the decisions from table 1. If Activity Type = Leisure non-sport, then PAE65a_leisure= mpae65; If Activity Type = Sports / exercise, then PAE65a_sports= mpae65; If Activity Type = Household, then PAE65a_household= mpae65; Note: Here we only classify them into three categories because there is no activity for 65a in the other two categories.
PAE65a_vig PAE65a_mod PAE65a_light	Recoding PAE65a into activity intensity levels	Here, we will categorize each element in PAE65a into 3 intensity levels (vigorous, moderate, light) See Table 1 for decisions of variables below and table 2 for hard-coding of variables based on the decisions from table 1. If Activity Intensity is not missing, then do; If Activity Intensity = Light, then PAE65a_light= mpae65; If Activity Intensity = Moderate, then PAE65a_mod= mpae65; If Activity Intensity = Vigorous, then PAE65a_vig= mpae65;

totpae69_min to totpae72_min	Total minutes for pae69- pae72	pae69a_min=pae69a * 60;
		totpae69_min=sum(pae69a_min, pae69b);
		
		pae70a_min=pae70a * 60;
		totpae70_min=sum(pae70a_min, pae70b);
		
		pae71a_min=pae71a * 60;
		totpae71_min=sum(pae71a_min, pae71b);
		
		pae72a_min=pae72a * 60;
		totpae72_min=sum(pae72a_min, pae72b);

References

1. Francis K, Feinstein R. A simple height-specific and rate-specific step test for children. South Med J. 1991;84:169-74. (in eng).
2. Eisenmann JC, Laurson KR, Welk GJ. Aerobic fitness percentiles for U.S. adolescents. Am J Prev Med. 2011;41:S106-10. (in eng).
3. Welk GJ, Laurson KR, Eisenmann JC, Cureton KJ. Development of youth aerobic-capacity standards using receiver operating characteristic curves. Am J Prev Med. 2011;41:S111-6. (in eng).
4. McMurray RG, Bangdiwala SI, Harrell JS, Amorim LD. Adolescents with metabolic syndrome have a history of low aerobic fitness and physical activity levels. Dyn Med. 2008;7:5. (in eng).
5. Ruiz JR, Ortega FB, Rizzo NS, Villa I, Hurtig-Wennlof A, Oja L, et al. High cardiovascular fitness is associated with low metabolic risk score in children: the European Youth Heart Study. Pediatr Res. 2007;61:350-5. (in eng).

13.1 PAE_TRANSPORT: Transportation activity, times/mo

This is a numeric variable that determines the times/month a child participates in transportation activities.

PAE_TRANSPORT =SUM(MPAE1, MPAE2, MPAE3)

Source variables: MPAE1, MPAE2, MPAE3

13.2 PAE_SCHOOL: School activity, times/mo

This is a numeric variable that determines the times/month a child participates in school activities.

PAE_SCHOOL =SUM(MPAE4, MPAE5, MPAE6)

Source variables: MPAE4, MPAE5, MPAE6

13.3 PAE_LEISURENONSPORT: Leisure non-sport activity, times/mo

This is a numeric variable that determines the times/month a child participates in leisure sport activities.

PAE_LEISURENONSPORT=SUM(OF MPAE7-MPAE15, OF MPAE59, OF PAE65A_LEISURE)

Source variables: MPAE7-MPAE15, MPAE59, PAE65A_LEISURE

13.4 PAE_SPORT: Sports/exercise activity, times/mo

This is a numeric variable that determines the times/month a child participates in exercise sports activities.

PAE_SPORT =SUM(OF MPAE16- MPAE58, OF MPAE60- MPAE64, OF PAE65A_SPORTS)

Source variables: MPAE16- MPAE58, MPAE60- MPAE64, PAE65A_SPORTS

13.5 PAE_HOUSEHOLD: Household activity, times/mo

This is a numeric variable that determines the times/month a child participates in household activities. .

PAE_HOUSEHOLD =SUM(OF MPAE66-MPAE68, OF PAE65A_HOUSEHOLD)

Source variables: MPAE66-MPAE68, PAE65A_HOUSEHOLD

13.6 PAE_SCREEN_MIN: Screen activities (TV, video, computer, phone), min/day
This is a numeric variable that determines the times/month a child participates in transportation activities. The variable is set to missing if PAE_SCREEN_MIN > 1440.

PAE_SCREEN_MIN =SUM(TOTPAE69_MIN, TOTPAE70_MIN, TOTPAE71_MIN, TOTPAE72_MIN)

Source variables: TOTPAE69_MIN, TOTPAE70_MIN, TOTPAE71_MIN, TOTPAE72_MIN

13.7 PAE_SED: Sedentary behavior, times/mo
This is a numeric variable that determines the times/month a child participates in sedentary activities.

PAE_SED = SUM (MPAE1, MPAE6, MPAE9, MPAE10, MPAE11, MPAE15)

Source variables: MPAE1, MPAE6, MPAE9, MPAE10, MPAE11, MPAE15

13.8 PAE_LIGHT: Light activity, times/mo
This is a numeric variable that determines the times/month a child participates in light activities.

PAE_LIGHT = SUM (MPAE7, MPAE12, MPAE13, MPAE14, MPAE19, MPAE20, MPAE62, MPAE64, MPAE68, PAE65A_LIGHT)

Source variables: MPAE7, MPAE12, MPAE13, MPAE14, MPAE19, MPAE20, MPAE62, MPAE64, MPAE68, PAE65A_LIGHT

13.9 PAE_MOD: Moderate activity, times/mo
This is a numeric variable that determines the times/month a child participates in moderate activities.

PAE_MOD = SUM (MPAE2, MPAE4, MPAE5, MPAE8, MPAE17, MPAE18, MPAE21, MPAE22, MPAE24, MPAE26, MPAE27, MPAE30, MPAE32, MPAE34, MPAE37, MPAE38, MPAE39, MPAE41, MPAE48, MPAE51, MPAE52, MPAE54, MPAE57, MPAE59, MPAE61, MPAE63, MPAE66, MPAE67, PAE65A_MOD)

Source variables: MPAE2, MPAE4, MPAE5, MPAE8, MPAE17, MPAE18, MPAE21, MPAE22, MPAE24, MPAE26, MPAE27, MPAE30, MPAE32, MPAE34, MPAE37, MPAE38, MPAE39, MPAE41, MPAE48, MPAE51, MPAE52, MPAE54, MPAE57, MPAE59, MPAE61, MPAE63, MPAE66, MPAE67, PAE65A_MOD

13.10 PAE_VIG: Vigorous activity, times/mo
This is a numeric variable that determines the times/month a child participates in vigorous activities.

PAE_VIG=SUM(MPAE3, MPAE16, MPAE23, MPAE25, MPAE28, MPAE29, MPAE31, MPAE33, MPAE35, MPAE36, MPAE40, MPAE42, MPAE43, MPAE44, MPAE45, MPAE46, MPAE47, MPAE49, MPAE50, MPAE53, MPAE55, MPAE56, MPAE58, MPAE60, PAE65A_VIG)

Source variables: MPAE3, MPAE16, MPAE23, MPAE25, MPAE28, MPAE29, MPAE31, MPAE33, MPAE35, MPAE36, MPAE40, MPAE42, MPAE43, MPAE44, MPAE45, MPAE46, MPAE47, MPAE49, MPAE50, MPAE53, MPAE55, MPAE56, MPAE58, MPAE60, PAE65A_VIG

13.11 PAE_MVPA: Moderate or vigorous activity, times/mo

This is a numeric variable that determines the times/month a child participates in moderate or vigorous activities.

PAE_MVPA = SUM (PAE_VIG, PAE_MOD)

Source variables: PAE_VIG, PAE_MOD

14. Nutrition - 2010 Healthy Eating Index (2010-HEI)

14.1 HEI2010: Healthy Eating Index- 2010

The Healthy Eating Index-2010 (HEI-2010; Guenther et al., 2013) is a measure of overall diet quality, independent of quantity, which can be used to assess compliance with the *2010 Dietary Guidelines for Americans* and to monitor changes in dietary patterns. It includes twelve dietary components (nine adequacy and three moderation components) that reflect key aspects of diet quality, including fruit, vegetables, grains, dairy, protein foods, fatty acids, sodium, and empty calories. Table 1 lists the components, the optimal (maximum) number of points, and the criteria for assignment of the lowest and highest possible scores for each component. Components scores can range from 0-5, 0-10, or 0-20, and 2010-HEI score ranges from 0 to 100 with a higher score indicating greater consistency of the diet with the 2010 Dietary Guidelines for Americans. The performance of the HEI–2010 has been evaluated through an assessment of its content validity, construct validity, and reliability.

Table 1. Optimum score and standard for maximum and minimum scores for each 2010 HEI component

2010 HEI Component	Optimum score	Standard for maximum score (optimum)	Standard for minimum score of zero
Adequacy			
1. Total Fruit[a]	5	≥ 0.8 cup equiv/1,000 kcal	No fruit
2. Whole Fruit[b]	5	≥ 0.4 cup equiv /1,000 kcal	No whole fruit
3. Total Vegetables[c]	5	≥ 1.1 cup equiv/1,000 kcal	No vegetables
4. Greens and Beans[c]	5	≥ 0.2 cup equiv/1,000 kcal	No dark green vegetables, beans or peas
5. Whole Grains	10	≥ 1.5 oz equiv/1,000 kcal	No whole grains
6. Dairy[d]	10	≥ 1.3 cup equiv/1,000 kcal	No dairy
7. Total Protein Foods[e]	5	≥ 2.5 oz equiv/1,000 kcal	No protein foods
8. Seafood and Plant Proteins[ef]	5	≥ 0.8 oz equiv/1,000 kcal	No seafood or plant protein
9. Fatty Acids[g]	10	(PUFAs+MUFAs)/SFAs >2.5	(PUFAs+MUFAs)/SFAs ≤1.2
Moderation			
10. Refined Grains	10	≤ 1.8 oz equiv/1,000 kcal	≥ 4.3 oz equiv/ 1,000 kcal
11. Sodium	10	≤ 1.1 gram/1,000 kcal	≥ 2.0 gram / 1,000 kcal
12. Empty Calories[h]	20	≤ 19% of energy	≥ 50% of energy

a Includes 100% fruit juice.
b Includes all forms except fruit juice.
c Includes any beans and peas not counted as Total Protein Foods.
d Includes all milk products, such as fluid milk, yogurt, cheese, and fortified soy beverages.
e Beans and peas are included here (and not with vegetables) when the Total Protein Foods standard is otherwise not met.
f Includes seafood, nuts, seeds, soy products (other than beverages) as well as beans and peas counted as Total Protein Foods.
g Includes seafood, nuts, seeds, soy products (other than beverages) as well as beans and peas counted as Total Protein Foods.
h Calories from solid fats, alcohol, and added sugars; threshold for counting alcohol is >13 g/1000 kcal.

In SOL Youth, the HEI-2010 was calculated from the average of available 24hr dietary recall data. At least one dietary recall is required to calculate the HEI-2010 score. From the 1,466 SOL Youth children, 1,305 (90%) have both dietary recalls and only 13 children have none. Only 54 dietary recalls were excluded because the daily energy intake (variable Y04A20) was considered too extreme (defined as below the sequence-gender specific 1st percentile or above the 99th percentile). Children with at least one dietary recall have a HEI-2010 score; only 13 children have no dietary recalls and 1,305 (90%) children had it calculated using both recalls.

Table 2. Energy intake (kcal/day) percentiles 1 and 99 by dietary recall and sex, SOL Youth

Recall	Sex	N	Percentile 1	Percentile 99
1st	Girl	727	480.50	3558.54
2nd	Girl	700	441.85	3257.04
1st	Boy	701	519.36	4287.35
2nd	Boy	684	512.34	3764.39
Total		2,812		

We followed four general steps to calculate the HEI-2010:

1. Quantify each of the 12 components at the dietary recall level.
2. Average available dietary recalls.
3. Score each component
4. Compute the HEI-2010 score as the sum of scores for individual 12 components.

We used the "NDSR Guide to Create Variables Needed to Calculate Scores for Each Component of the HEI-2010" developed by the Nutrition Coordinating Center (NCC) at University of Minnesota, Minneapolis, MN. Table 3 has the specific NDSR food subgroups used to define each component of HEI-2010. Specifically, for 10 components we calculated the food groups in cup equivalents per day, and then calculated food group density by dividing by daily energy intake per 1,000kcal. Empty calories was expressed as a percentage of total energy intake, and fatty acids as a ratio of polynsaturated fatty acids (PUFAS) and monounsaturated (MUFAS) of total saturated fatty acids (SFAs). The scores for each individual component were computed according to the formulas given in Table 1. Intermediate intakes were scored proportionately between zero and the maximum score.

For each component (e.g. total fruits) we are releasing its score (e.g. HEI1) plus intermediate variables used to calculate it (at recall level, averaged, cups equivalents, cups equivalents/1,000 kcal).

Units	1st recall	2nd recall	Average
cup equivalents per day	TOT_FRUIT_R1	TOT_FRUIT_R2	TOT_FRUIT_AVG
cup equivalents/1,000 kcal per day	TOT_FRUIT_DEN_R1	TOT_FRUIT_DEN_R2	TOT_FRUIT_DEN_AVG

References

- Guenther PM, Kirkpatrick SI, Reedy J, Krebs-Smith SM, Buckman DW, Dodd KW, Casavale KO, Carroll RJ. The Healthy Eating Index-2010 is a valid and reliable measure of diet quality according to the 2010 Dietary Guidelines for Americans. J Nutr. 2014 Mar;144(3):399-407.
- Guenther PM, Casavale KO, Reedy J, Kirkpatrick SI, Hiza HA, Kuczynski KJ, Kahle LL, Krebs-Smith SM. Update of the Healthy Eating Index: HEI-2010. J Acad Nutr Diet. 2013 Apr;113(4):569-80.
- NDSR Guide to Creating Variables Needed to Calculate Scores for Each Component of the Healthy Eating Index-2010 (HEI-2010) developed by the Nutrition Coordinating Center (NCC), University of Minnesota, Minneapolis, MN

Table 3. Definition of HEI-2010 components using HCHS/SOL variables from 24hr dietary recalls.		
Variable	Label	Definition
TOT_PROTEIN_V1 (V1 stands for version 1, without legumes)	Total Protein Foods (ounce equivalents) without legumes	sum (MEAT_RED, MEAT_LUNCH, MEAT_POUL, MEAT_FISH, MEAT_ORG, MEAT_EGG, MEAT_NUT, Y09A77, Y09A83)
TOT_PROTEIN_V2 (V2 stands for version 2, without legumes)	Total protein foods (ounce equivalents) with legumes	sum(MEAT_RED, MEAT_LUNCH, MEAT_POUL, MEAT_FISH, MEAT_ORG, MEAT_EGG, MEAT_NUT, Y09A77, Y09A83, [Y09A17*2]]
TOT_PROTEIN_STANDARD	Total protein foods standard is met	If (TOT_PROTEIN_V1 / (Y04A20/1000))>=2.5 then TOT_PROTEIN_STANDARD = 1 ELSE TOT_PROTEIN_STANDARD = 0
ALLOWABLE_ALCOHOL	Allowable alcohol for Empty Calories (g)	Y04A20*0.013
EXCESS_ALCOHOL	Excess alcohol for Empty Calories (kcal)	If .Z<Y04A26 <= ALLOWABLE_ALCOHOL then EXCESS_ALCOHOL=0 If Y04A26 > ALLOWABLE_ALCOHOL then EXCESS_ALCOHOL=((Y04A26 – ALLOWABLE_ALCOHOL)*7)
TOT_FRUIT	Total fruit (cup equivalents)	FRUIT_ALL / 2
WHOLE_FRUIT	Whole fruit (cup equivalents)	sum(Y09A6, Y09A7, Y09A8, Y09A9, Y09A10)/2
TOT_VEG	Total vegetables (cup equivalents)	sum(Y09A11, Y09A12, Y09A13, Y09A14, Y09A15, Y09A16, Y09A17*(TOT_PROTEIN_STANDARD) + Y09A18, Y09A19, Y09A20) /2
GREENS_BEAN	Greens and beans (cup equivalents)	[Y09A11 + Y09A17*(TOT_PROTEIN_STANDARD)]/2
WHOLE_GRAINS	Whole grains (ounce equivalents)	sum[Y09A22 , Y09A25 , Y09A28 , Y09A31 , Y09A34 , Y09A37, Y09A40, Y09A43, Y09A46, Y09A49, Y09A52, Y09A53]
DAIRY	Dairy (cup equivalents)	sum[Y09A84, Y09A85, Y09A86, Y09A87, Y09A88, Y09A89, Y09A90, Y09A91 ,Y09A92, Y09A95, Y09A96, Y09A97, Y09A98, Y09A99, Y09A100, Y09A101, Y09A102, Y09A103, Y09A104, Y09A105, [Y09A106/3], Y09A108 , Y09A109, Y09A114, Y09A115]
TOT_PROTEIN	Total Protein Foods (ounce equivalents)	If TOT_PROTEIN_STANDARD = 1 then TOT_PROTEIN=TOT_PROTEIN_V1 Else TOT_PROTEIN=TOT_PROTEIN_V2
SEA_PLANT_PROTEIN	Seafood and Plant Proteins (ounce equivalents)	sum(MEAT_FISH ,MEAT_NUT, Y09A83, [Y09A17* 2]*(1-TOT_PROTEIN_STANDARD)
FATTY_ACIDS	Fatty acids (ratio)	sum(Y04A30, Y04A29) / Y04A28
REFINED_GRAINS	Refined grains (ounce equivalents)	sum [Y09A23, Y09A24, Y09A26, Y09A27, Y09A29, Y09A30, Y09A32, Y09A33, Y09A35,Y09A36, Y09A38, Y09A39, Y09A41, Y09A42, Y09A44, Y09A45, Y09A47, Y09A48, Y09A50, Y09A51, Y09A54]
SODIUM	Sodium (grams)	Y04A67/1000
EMPTY_CALS	Empty calories (from solid fats, alcohol above>13 g/1,000 kcal, and added sugars)	sum((Y04A28*9), (Y04A132*9), (Y04A184*4), EXCESS_ALCOHOL
EMPTY_CALS_PCT	Empty calories percent of total calories	(EMPTY_CALS/Y04A20)*100
TOT_FRUIT_DEN Similar density variables for the other components except for FATTY_ACIDS and EMPTY_CALORIES	Total fruit density (cup equivalents per 1000 kcal)	TOT_FRUIT / (Y04A20/1000)

15.Variable Modifications

Data release	Variable	# participants changed	Notes
5/2015	PREDIABETES	0	Updated to reflect HbA1c cutoff change to 5.7 (from 6.0)
8/2015	HOMA_IR	1382	Calibrate insulin variable LABA11 units from pmol/L to mU/mL in the formula (divide by 6).
8/2015	INSULIN_RESIST	571	Uses updated, calibrated HOMA_IR variable.
2/2015	DIABETES_GLUCOSE	0	Updated cutoff to be >= 126 instead of >126 in previous version
6/2016	STRAT	9	Updated to collapse low sample size stratum into similar stratum in the same field center as follows: STRAT= 28 -> changed to STRAT=13 STRAT= 18 -> changed to STRAT=16
6/2016	FPG_FLAG	29	Updated cutoff to be >= 100 instead of >100 in previous version.
6/2016	HBA1C_FLAG	59	Updated cutoff to be >= 5.7 instead of >5.7 in previous version.
6/2016	INSULIN_RESIST	0	Updated cutoff to be >= 2.5 instead of >2.5 in previous version.
6/2016	MET_SYNDROME_ATP	4	Updated due to changes with FPG_FLAG
6/2016	MET_SYNDROME_IDF	2	Updated due to changes with FPG_FLAG
6/2016	CVD_RISK_COUNT	84	Updated due to changes with FPG_FLAG and HBA1C_FLAG
6/2016	CVD_RISK_C4	62	Updated due to changes with FPG_FLAG and HBA1C_FLAG
6/2016	CVD_RISK_C3	35	Updated due to changes with FPG_FLAG and HBA1C_FLAG
6/2016	CVD_RISK_BIN	22	Updated due to changes with FPG_FLAG and HBA1C_FLAG
6/2016	WHOLE_GRAINS_R1	93	Updated to correct coding typo by replacing Y09A41(ready-to-eat cereal from refined grains) with Y09A43 (cake, cookies, pies, pastries, Danish, doughnuts and cobblers from whole grains)
6/2016	WHOLE_GRAINS_R2	87	Updated to correct coding typo by replacing Y09A41(ready-to-eat cereal from refined grains) with Y09A43 (cake, cookies, pies, pastries, Danish, doughnuts and cobblers from whole grains)
6/2016	WHOLE_GRAINS_AVG	158	Updated due to changes with WHOLE_GRAINS_R1 and WHOLE_GRAINS_R2
6/2016	WHOLE_GRAINS_DEN_R1	93	Updated due to changes with WHOLE_GRAINS_R1 and WHOLE_GRAINS_R2

6/2016	WHOLE_GRAINS_DEN_R2	87	Updated due to changes with WHOLE_GRAINS_R1 and WHOLE_GRAINS_R2
6/2016	WHOLE_GRAINS_DEN_AVG	158	Updated due to changes with WHOLE_GRAINS_R1 and WHOLE_GRAINS_R2
6/2016	HEI5	132	Updated due to changes with WHOLE_GRAINS_R1 and WHOLE_GRAINS_R2
6/2016	HEI2010	132	Updated due to changes with WHOLE_GRAINS_R1 and WHOLE_GRAINS_R2