



DERIVED VARIABLE DICTIONARY

January 2020

**Prepared by the Collaborative Studies Coordinating Center
UNC Chapel Hill, Department of Biostatistics**

Version 4.23 – Investigator Use Data Sets (_INV4 data file)

HCHS/SOL Derived Variable Dictionary

Table of Contents

SUMMARY OF UPDATES IN JANUARY 2020 DATA RELEASE	10
1. DESIGN	11
1.1 Cohort (Yearly Recruitment Assignment from Sample Address List).....	11
1.2 PSU_ID (Primary Sampling Unit ID Clustering Variable- masked).....	11
1.3 STRAT (Sampling Design Stratification Variable - masked)	11
1.4 LISTNUM (List Address Record Assigned).....	11
1.5 PSU_CNT (BG-level Hispanic Household Frame Counts).....	11
1.6 STRAT_CNT (Stratum-level Hispanic Household Frame Count).....	11
1.7 LIST_CNT (Sample List - Level Hispanic).....	11
1.8 Weight_Final_Expanded (Expanded Census 2010 Calibrated, Trimmed, Nonresponse Adjusted Weights - Year 1-3 (frozen as of 4/25/2012))	12
1.9 Weight_Final_Norm_Overall (Overall Normalized Census 2010 Calibrated Trimmed NonResponse Adjusted Final Weights - Year 1-3 (frozen as of 11/30/2012))	12
1.10 Weight_Final_Norm_Center (FC-specific Normalized Census 2010 Calibrated Trimmed NonResponse Adjusted Final Weights - Year 1-3 (frozen as of 4/25/2012))..	12
2. ADMINISTRATIVE	13
2.1 Center (Participant’s field center).....	13
2.2 Centernum – Numeric (Participant’s field center).....	13
2.3 Clindate (Date of the participant’s clinic visit).....	13
2.4 Consent (Informed consent status for participation in study)	13
2.5 Full_AFU_Eligible (Core Baseline Requirements for AFU Eligibility)	14
2.6 Genotyping (Full consent for genotyping)	14
2.7 DNA_Available (DNA Available (1=Yes,0=No))	15
2.8 NP_Permit (Permission for non-profit use of data).....	16
2.9 External_Permit (Permission for investigators outside HCHS/SOL)	16
2.10 Commercial_Permit (Permission for for-profit use of data)	16
2.11 Any_Permit (Any permission for public use data)	17
3. SOCIO-DEMOGRAPHIC	18
3.1 Age (Age in years at the time of participant’s clinic visit)	18
3.2 Agegroup_C2 (2-level grouped age of participant)	18
3.3 Agegroup_C3 (3-level grouped age of participant)	18
3.4 Agegroup_C6 (6-level grouped age of participant)	19
3.5 Agegroup_C6_NHANES (6-level grouped age – NHANES standardization) ...	19
3.6 Gender (Gender (F=Female, M=Male))	20
3.7 Gendernum (Gender – Numeric (0=Female, 1=Male))	20
3.8 Education_C2 (2-level group education level).....	20
3.9 Education_C3 (3-level group education level).....	21
3.10 Education_C4 (4-level group education level).....	22
3.11 Income (Grouped yearly household income)	23

3.12	Income_C3 (3-level grouped yearly household income)	24
3.13	Income_C5 (5-level grouped yearly household income)	24
3.14	POVPCT (Household income, percent of poverty)	24
3.15	POVPCT_C4 (3-level grouped Household income, percent of poverty)	26
3.16	Marital_Status (4-level grouped marital status)	26
3.17	Employed (4-level grouped employment status)	27
3.18	N_HC (Health Insurance Coverage - Current)	27
3.19	Occupation_Curr (Grouped current occupation for employed participants)	28
3.20	Occupation_Long (Grouped longest held occupation)	28

4. ACCULTURATION 29

4.1	Bkgrd1 (Re-classification of Hispanic/Latino Background)	29
4.2	Bkgrd1_C7 (7-level re-classification of Hispanic/Latino Background)	30
4.3	Bkgrd1_C6 (6-level re-classification of Hispanic/Latino Background)	30
4.4	Site_Bkgrd (Cross-Classification of Center and Hispanic/Latino Background)	31
4.5	Site_Bkgrd_45Plus (Cross-Classification of Center and Hispanic/Latino Background, 45+ Years Only)	32
4.6	Race (Race Group (self-report))	33
4.7	US_Born (Grouped place of birth – 50 US States only)	33
4.8	US_Native (Grouped place of birth – US states and territories)	34
4.9	Lang_Pref (Language Preference for Baseline Examination)	34
4.10	IMGEN_C2 (2-level Immigrant Generation)	34
4.11	IMGEN_C4 (4-level Immigrant Generation)	35
4.12	YRSUS (Years lived in the US)	36
4.13	YRSUS_C2 (2-level grouped years lived in the US (50 States))	36
4.14	YRSUS_C3 (3-level grouped years lived in the US (50 States))	36
4.15	SASH_ALL (Short acculturation scale for Hispanic)	37
4.16	SASH_LANG (Short acculturation scale for Hispanics – language subscale)	37
4.17	SASH_MEDIA (Short acculturation scale for Hispanics – media)	37
4.18	SASH_SOC (Short acculturation scale for Hispanics – Social subscale)	37
4.19	ETHIS (Ethnic Identification Score)	38
4.20	Language_Subscore_MESA (Language Subscore - MESA)	38
4.21	Nativity_Subscore_MESA (Nativity Subscore - MESA)	39
4.22	Accult_MESA (Acculturation Score - MESA)	39
4.23	Age_IMMI (Age of immigration among participants who were not born in US mainland (50 US States + DC))	40

5. ANTHROPOMETRY 41

5.1	BMI (Body Mass Index kg/m ²)	41
5.2	BMIGRP_C4 (4-level grouped Body Mass Index - WHO)	41
5.3	BMIGRP_C6 (6-level grouped Body Mass Index - WHO)	41
5.4	Waist_hip (Waist to Hip ratio)	42
5.5	Height (Height corrected in cm)	42

6. RESPIRATORY 43

6.1	Asthma_Ever (Ever had asthma)	43
6.2	Asthma_Ever_MD (Ever had asthma with MD diagnosis)	43

6.3	Asthma_Curr (Currently have asthma)	43
6.4	Asthma_Curr_MD (Currently have asthma with ever MD diagnosis)	44
6.5	Asthma_C4 (4-level grouped Asthma Status)	44
6.6	Asthma_C4_MD (4-level grouped Asthma Status – by MD diagnosis)	44
6.7	COPD_Self (Self-report of ever had COPD/Emph).....	45
6.8	COPD_EVER (Self-report of ever had COPD/Emph or CB)	45
6.9	MRC_CB (Chronic Productive Cough 3+ mths in 2+ yrs)	45
6.10	Wheeze_SB (Wheezing ever with shortness of breath).....	46
6.11	Wheeze_Last12 (Wheezing in last 12 months).....	46
6.12	Phlegm (Phlegm day/night 3 months in a row).....	46
6.13	Pneumonia (Pneumonia - by MD diagnosis).....	47
6.14	Valid_spirometry (Valid spirometry studies, pre-BD).....	47
6.15	Valid_spirometry_postBD (Valid spirometry studies, post-BD)	47
6.16	Airflow_Obstruction (Obstructive Lung Disease).....	47
6.17	Airflow_Limitation (Airflow Limitation (pre FEV1/FVC<70 only))	48
6.18	COPD_by_BD (COPD by bronchodilator test)	48
6.19	COPD_by_BD_AllPre (COPD by BD test (for all valid pre_BD)).....	48
6.20	BD_response_abs (Bronchodilator Responsive by 15% change).....	49
6.21	FEV1_FVC_Ratio (FEV ₁ to FVC %).....	49
6.22	POSTBD_PREBD_DIFF (Post(FEV1%pred)-Pre(FEV1%pred))	49
6.23	POSTBD_PREBD_DIFF_GT12 (Relative Bronchodilator Responsive by 12% change) 50	
6.24	POSTBD_PREBD_DIFF_GT15 (Relative Bronchodilator Responsive by 15% change) 50	
6.26	POSTBD_PREBD_DIFF_GT12_2 (Relative Bronchodilator Responsive by 12% change (among pre FEV1/FVC<70 only))	51
6.27	POSTBD_EXPECTED (Post-BD Expected (among valid spirometry))	51
6.28	VALID_SPIROMETRY_POSTBD2 (Valid post-spirometry (A-D grading) adjusted for unexpected BDs).....	51
7.	WELL-BEING/QUALITY OF LIFE	53
7.1	CESD10 (10-item CES-D summary score)	53
7.2	STAI10 (10-item State-Trait Anxiety Inventory summary score)	53
8.	SF-12V2 HEALTH SURVEY	55
8.1	Agg_ment (Aggregate Mental Health Scale).....	55
8.2	Agg_phys (Aggregate Physical Health Scale).....	55
9.	BLOOD PRESSURE MEASURES	56
9.1	Left_ABI, Right_ABI, and ABI_Overall (Ankle Brachial Index –occ. failure included).....	56
9.2	LEFT_ABIGRP_C4, RIGHT_ABIGRP_C4 (4-level grouped Ankle Brachial Index (occ. failure incl.))	57
9.3	ABIGRP_C4 (4-level grouped Ankle Brachial Index (occ. failure incl.))	57
9.4	Left_ABI2, Right_ABI2, and ABI2_Overall (Ankle Brachial Index - version 2)..	57
9.5	ABIGRP2_C4 (4-level grouped Ankle Brachial Index - version 2).....	58

10.	LABORATORY MEASURES	59
10.1	CKD (Chronic Kidney Disease using eGFR only - NIDDK)	59
10.2	Insulin_Fast (Insulin, fasting (calibrated, converted to mU/L))	59
10.3	Insulin_OGTT (Insulin, post OGTT (calibrated, converted to mU/L))	60
10.4	Insulin_Date (Date of Insulin Measures - All)	60
10.5	Dyslipidemia (High LDL, low HDL, or high triglycerides - NIH)	60
10.6	Dyslipidemia_C3 (3-Category Dyslipidemia)	61
10.7	Cardiac_Risk_Ratio (Cardiac Risk Ratio (TC/HDL) - AHA)	61
10.8	Dyslipidemia_TCHDL (Dyslipidemia based on Cardiac Risk Ratio > 5)	61
10.9	DYS_TCHDL_MED (Dyslipidemia based on Cardiac Risk Ratio > 5 or Lipid Lower Drug scanned at baseline)	61
10.10	High_total_Chol (Hypercholesterolemia (from LABA66, LABA68, LABA69, MED_LLD))	62
10.11	TOTAL_CHOL_C3 (3-Category Total Cholesterol)	62
10.12	Fasttime (Fasting time in hours)	63
10.13	IFG (Impaired Fasting Glucose - AHA)	63
10.14	IGT (Impaired Glucose Tolerance - AHA)	63
10.15	GFR (Glomerular Filtration Rate, estimated)	63
10.16	GFRscys (eGFR based on serum cystatin C w/o demographics)	64
10.17	GFRscys_scr (eGFR based on serum cystatin C, serum creatinine, gender, age and race)	64
10.18	GFRscr (eGFR based on serum creatinine, gender, age, and race)	64
11.	GLOBAL PHYSICAL ACTIVITY QUESTIONNAIRE (GPAQ)	66
11.1	GPAQ_rec_mod (Recreational physical (moderate) activity (min/day))	66
11.2	GPAQ_rec_vig (Recreational physical (vigorous) activity (min/day))	66
11.3	GPAQ_rec (Recreational physical activity (min/day))	67
11.4	GPAQ_work_mod (Work-related (moderate) physical activity (min/day))	67
11.5	GPAQ_work_vig (Work-related (vigorous) physical activity (min/day))	67
11.6	GPAQ_work (Work-related physical activity (min/day))	68
11.7	GPAQ_total_mod (Total physical (moderate) activity (min/day))	68
11.8	GPAQ_total_vig (Total physical (vigorous) activity (min/day))	68
11.9	GPAQ_trsport (Transportation-related physical activity (min/day))	69
11.10	GPAQ_total (Total physical activity (min/day))	69
11.11	GPAQ_total_MET (Total physical activity (MET-min/day))	69
11.12	GPAQ_MOD_WEEK (GPAQ - Total physical (moderate) activity (min/week))	70
11.13	GPAQ_VIG_WEEK (GPAQ - Total physical (vigorous) activity (min/week))	70
11.15	GPAQ_level (Physical activity level)	70
11.16	GPAQ_PAG2008 (Activity level per 2008 PA guidelines (categorical - 4 levels))	71
11.17	GPAQ_PAG2008YN (Meets 2008 activity level guidelines (1=Yes, 0=No))	72
12.	CLINICAL CHARACTERISTICS	73
12.1	Hypertension (Hypertension (BP \geq 140/90 and Med Use))	73
12.2	Hypertension2 (Hypertension using NHANES definition)	73
12.3	Hypertension_C4 (4-level grouped hypertension)	73
12.4	Hypertension2_AHA (Hypertension using new ACC/AHA Guidelines definition)	74

12.5	Hypertension_AHA_C5 (5-level grouped hypertension using new ACC/AHA Guidelines definition).....	74
12.6	Hypert_Awareness (Awareness of Hypertension - NHANES).....	75
12.7	Hypert_Treatment (Treatment of Hypertension - NHANES)	75
12.8	Hypert_Control (Control of Hypertension - NHANES).....	76
12.9	HOMA_IR (HOMA index of Insulin Resistance)	76
12.10	HOMA_B (HOMA index of Beta Cell Function)	76
12.11	Diabetes_self (Diabetes – self-report only)	76
12.12	Diabetes_Lab (3-level grouped Diabetes – Lab ADA Guideline)	77
12.13	SelfMed_Antidiab (Anti-Diabetics Medication used – Self-reported).....	77
12.14	Diabetes1 (3-level grouped Diabetes ver.1).....	78
12.15	Diabetes2 (3-level grouped Diabetes – ADA Guideline plus scanned Med)	78
12.16	Diabetes2_Indicator (Diabetes 2 Indicator (ADA))	79
12.17	Diabetes3 (3-level grouped Diabetes – ADA Guideline plus Self-reported Diabetes).....	79
12.18	Diabetes_C4 (4-level grouped diabetes (none/pre/treated/untreated))	80
12.19	Diab_Diag (Diabetes Diagnosis)	80
12.20	Diab_FamHist (Family History of Diabetes)	81
12.21	DM_Aware (Diabetes Awareness at Baseline)	81
12.22	DM_Control (Diabetes controlled: DM Classified at Baseline and A1C < 7%)	81
12.23	Diabetes4 (3-level grouped Diabetes – ADA Guideline plus self-reported Med)	82
12.24	Diabetes5 (3-level grouped Diabetes – ADA Guideline plus Self-reported diabetes diagnosis or Self-reported Med)	82
12.25	PreCHD_No_Angina (Prevalent Cardiovascular Heart Disease)	83
12.26	PreCHD_Angina (Prevalent Cardiovascular Heart Disease, including angina)	83
12.27	Ever_Angina_Relative (At least one relative had report of angina).....	83
12.28	Ever_MI_Relative (At least one relative had report of Heart Attack)	84
12.29	Ever_CABG_Relative (At least one relative had report of CABG or angioplasty)	84
12.30	FH_CHD (Family history of CHD including self-reported angina)	85
12.31	CHD_SELF (Self-reported CHD)	85
12.32	CDCR (Cerebrovascular Disease or Carotid Revascularization, self-report) ...	86
12.33	METS_NCEP (Metabolic Syndrome – NCEP-ATP III)	86
12.34	METS_NCEP2 (Metabolic Syndrome w/ self-reported meds – NCEP-ATP III)	87
12.35	METS_IDF (Metabolic Syndrome – IDF).....	87
12.36	MI_ECG (MI defined by ECG results)	89
12.37	Stroke (Prevalent Stroke, self-report).....	89
12.38	Stroke_TIA (Prevalent Stroke or TIA, self-report)	89
12.39	FH_Stroke (Family History of Stroke)	89
12.40	PRECVD (Prevalent CVD (CHD or Stroke, not counting Angina or TIA))	90
12.41	Cigarette_Use (1=Never,2=Former,3=Current).....	90
12.42	Current_Smoker (Current Smoker (from CIGARETTE_USE)).....	91
12.43	Exposure_Year (Adjusted Years Smoked).....	91
12.44	Cigarettes_Year (Adjusted Cigarette Years).....	92
12.45	Cigarette_Pack_Years (Cigarette Pack Years)	92
12.46	Alcohol_Use (1=Never,2=Former,3=Current)	92
12.47	TotalDrinks_Per_Week (Total Weekly Alcohol Consumption)	93
12.48	Alcohol_Use_Disorder (Alcohol Use Disorder Risk)	93
12.49	Alcohol_Use_Level (Alcohol Use Disorder Risk - NIAAA).....	93

12.50	Doctor_Visit (Doctor/HCP number of visits in past 12).....	94
12.51	Claudication_Int (Intermittent Claudication - either leg)	94
12.52	CVD_Frame (Framingham CVD composite criterion)	95
12.53	Frame_CVD_Risk_10Yr (Framingham CVD - 10 year risk (w/lab predictors)).	96
12.54	HBA1C_SI (Glycosylated Hemoglobin in SI units (mmol/mol))	99
12.55	CIGARETTE_PACK_YEARS_C3 (Cigarette Pack Years (3-level, cutoff =10))	99

13. MEDICATIONS – SCANNED/TRANSCRIBED ONLY 100

13.1	Med_Current_HRT (Any Hormone Repl. Therapy)	101
13.2	Med_Estrogens (Estrogens)	101
13.3	Med_Estrogen_Combo (Estrogens – Combinations Only)	101
13.4	Med_Contraceptives (Contraceptives).....	101
13.5	Med_Progestins (Progestins).....	101
13.6	Med_SERM (Selective Estrogen Receptor Modulators)	102
13.7	Med_Fertility_Reg (Fertility Regulators).....	102
13.8	Med_LHRH (LHRH/GnRH Only).....	102
13.9	Med_LHRH_AG (LHRH/GnRH Agonist Analog)	102
13.10	Med_LHRH_ANT (LHRH/GnRH Antagonists)	102
13.11	Med_Antianginal (Antianginals)	102
13.12	Med_Antiarrhythmics (Antiarrhythmics)	103
13.13	Med_Antiasthmatics (Antiasthmatic or bronchodilator agents)	103
13.14	Med_Anticoag (Anticoagulants)	103
13.15	Med_Anticoag_Coumarin (Coumarin anticoagulants).....	103
13.16	Med_Anticoag_Heparin (Heparin anticoagulants).....	103
13.17	Med_Antidiab (Antidiabetics)	103
13.18	Med_Antihypert (Antihypertensives)	103
13.19	Med_Antihypert_ACEI (ACE Inhibitors)	103
13.20	Med_Antihypert_AT2RAs (Angiotension II Receptor Antagonists)	104
13.21	Med_Antiplatelet (Antiplatelets)	104
13.22	Med_Aspirin (Aspirin-containing analgesics)	104
13.23	Med_BB (Beta Blockers).....	104
13.24	Med_BB_Ophthalmic (Ophthalmic Beta Blockers).....	105
13.25	Med_CACB (Calcium Channel Blockers).....	105
13.26	Med_Cardiacycosides (Cardiac Glycosides).....	105
13.27	Med_Chemo (Chemotherapy).....	105
13.28	Med_Clopid (Clopidogrel (Plavix))	105
13.29	Med_COX2 (COX-2 inhibitors).....	105
13.30	Med_Diuretic (Diuretics).....	106
13.31	Med_Diuretic_Thiazide (Thiazide Diuretics).....	106
13.32	Med_Fibares_Nicoacid (Fibric/Nicotinic Acids (trt of TG and HDL))	106
13.33	Med_Insulin (Insulins)	106
13.34	Med_LLD (Lipid lowering drugs/Antihyperlipidemics (LLD)).....	106
13.35	Med_Metformin (Metformins)	107
13.36	Med_NSAID (NSAIDs).....	107
13.37	Med_OI_Steroid (Oral/inhalable Glucocorticosteroids)	107
13.38	Med_Statin (Statins)	107
13.39	Med_ANTIAXI (Antianxiety agents).....	107
13.40	Med_ANTIDEPRESS (Antidepressants).....	108

13.41	Med_ANTIDEPRE_SSRI (Antidepressants - selective serotonin re-uptake inhibitors).....	108
13.42	Med_ANTIDEPRE_TRICYC (Antidepressants - tricyclic antidepressants)	108
13.43	Med_ANTIPSYCHO (Antipsychotics)	108
13.44	Med_CORTICOSTEROIDS_INHALED (Inhaled Corticosteroid medication) .	108
13.45	Med_PDE4 (Phosphodiesterase IV Inhibitor Medication)	108
13.46	Med_BRONCHODILATOR (Bronchodilator and Antiasthmatic Agents medication) 108	
13.47	Med_PULMONARY (Pulmonary medication).....	109
14.	MEDICATIONS – OTHER	111
14.1	Asthma_Meds (Asthma Med Use – self/scanned (curr asthma by MD diag))	111
14.2	COPD_Meds (COPD Med Use – self/scanned (ever COPD by MD diag))	111
15.	NEUROCOGNITIVE	112
15.1	ECG_MAJ (Major ECG Abnormalities)	112
15.2	ECG_MIN (Minor ECG Abnormalities).....	112
15.3	No_Disability_6Item (No Disability on NEEA5-7 (1=Yes,0=No)).....	114
15.4	SEVLT_3Trials (SEVLT, Sum of 3 Trials (NEEA8=10)).....	114
15.5	SEVLT_Recall (SEVLT, Delay Recall (NEEA11 with re-coding)).....	114
15.6	Total_6Item (Total Score on 6 Item Screener (with re-coding))	114
15.7	WORDFREQ (Word Frequency (A+F) (NEEA12-13)).....	116
15.8	DIGITSYMBOL (Digit Symbol Substitution Score(NEEA14)).....	116
16.	DENTAL	117
16.1	Periodontal_Disease (Periodontal Disease - NIDCR)	117
16.2	PERIO_CDC_AAP (Periodontitis Severity using CDC/AAP Grades).....	118
16.3	PERIO_Extent (Percentage Interproximal Sites Positive for Periodontitis)	119
17.	SLEEP	121
17.1	ESS (Epworth Sleepiness Scale).....	121
17.2	ESS_GT10 (Excessive sleepiness (ESS > 10)).....	121
17.3	AHI_GE15 (Sleep Apnea defined as Apnea/Hypopnea Index (3% desat) ≥ 15)	121
17.4	SBD5 (Sleep Apnea Syndrome (Apnea/Hypopnea Index (3% desat) >= 5 & ESS >=10))	122
17.5	SLPDUR_WKDAY (Weekday sleep duration (hours))	122
17.6	SLPDUR_WKEND (Weekend sleep duration (hours)).....	122
17.7	SLPDUR (Average sleep duration (hours))	123
17.8	MID_SLEEP_TIME_WKDAY (Midpoint sleep time in clock hours; weekdays)	123
17.9	MID_SLEEP_TIME_WKEND (Midpoint sleep time in clock hours; weekend)	123
17.10	WHIIRS (Women’s Health Initiative Insomnia Rating Scale).....	124
18.	DIET	125
18.1	Diet_Score_JAMA (Diet Score from 1st JAMA paper)	125
18.2	Diet_Score_JAMA_C2 (Categorized Diet Score JAMA (1=Below, 2=Above 60th %-tile))	126

18.3	AHEI2010 (Alternative Healthy Eating Index 2010)	126
18.4	HEI2010 (Healthy Eating Index 2010).....	127
18.5	DASH (Version 1 – 2003 Eating Plan)	128
18.6	DASH (Version 2 – 2018 Eating Plan)	129
18.7	SUPPL_USE (Supplement use in the last 30 days).....	130
VARIABLE MODIFICATIONS		131
APPENDIX		135

Summary of updates in January 2020 data release

New Derived Variables

- AGEGROUP_C3
- EDUCATION_C4
- YRSUS_C3
- ABI2_OVERALL, LEFT_ABI2, RIGHT_ABI2
- PRECVD
- HYPERTENSION2_AHA - New ACC/AHA Guideline: 130/80 mmHg
- HYPERTENSION_AHA_C5 - New ACC/AHA Guideline: 130/80 mmHg
- DIABETES4 – ADA Lab criteria plus self-reported meds
- DIABETES5 – ADA Lab criteria plus self-reported diagnosis plus self-reported meds
- SUPPL_USE – Supplement use from NDSR 30-day supplement module
- HEI2010 – Healthy Index Score 2010
- MID_SLEEP_TIME_WKDAY and MID_SLEEP_TIME_WEND
- DASH2003 – 2003 Eating Plan
- DASH2018 – 2018 Eating Plan
- SELFMED_ANTIDIAB

Intermediate Variables Kept

- FASTTIME
- ELEVATED_BP_PLUSMEDS, ELEVATED_BP_SELFMEDS
- ABDOMINAL_OBESITY_IDF, ABDOMINAL_OBESITY_NCEP
- HIGH_TRIG, HIGH_TRIG_PLUSMEDS, LOW_HDL, LOW_HDL_PLUSMEDS
- IFG_IDF, IFG_NCEP
- GFR

Corrected Variables

- ESS_GT10. Sleepiness. Definition corrected from \geq to $>$
- PERIODONTAL_DISEASE. Algorithm corrected
- PERIO_EXTENT. Algorithm corrected
- METS_NCEP2. Gender source variable updated
- DM_AWARE. Algorithm corrected
- MED_STATIN. Two codes added
- MED_CACB. One code added
- MED_BB. One code added

Variable Name Changed. We removed “_V2” from some variable names because it denoted version 2 of a variable but the same extension is now used to name variables from Visit 2. This file only has BASELINE data.

- ABIGRP_C4_V2 → ABIGRP2_C4
- LEFT_ABI_V2 → LEFT_ABI2
- RIGHT_ABI_V2 → RIGHT_ABI2
- ABI_OVERALL_V2 → ABI2_OVERALL
- POSTBD_PREBD_DIFF_GT12_V2 → POSTBD_PREBD_DIFF_GT12_2
- VALID_SPIROMETRY_POSTBD_V2 → VALID_SPIROMETRY_POSTBD2

1. DESIGN

1.1 Cohort (Yearly Recruitment Assignment from Sample Address List)

Indicator of household sample address allocation and distribution for HCHS/SOL recruitment.

Cohort= 1, 2, 3 (year 1, year 2, or year 3 of the sample address recruitment cohorts)

1.2 PSU_ID (Primary Sampling Unit ID Clustering Variable- masked)

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.

1.3 STRAT (Sampling Design Stratification Variable - masked)

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)

1.4 LISTNUM (List Address Record Assigned)

Sample design variable used in statistical analyses that is the following:

Hispanic Surname versus non-Hispanic surname address list identifier from the original Genesys sampling address lists used in recruitment.

Response format: 1 = Non-Hispanic Surname Address List
2 = Hispanic Surname Address List

1.5 PSU_CNT (BG-level Hispanic Household Frame Counts)

Sample design clustering household frame count variable used in statistical analyses.

1.6 STRAT_CNT (Stratum-level Hispanic Household Frame Count)

Sample design stratification household frame variable used in statistical analyses.

1.7 LIST_CNT (Sample List - Level Hispanic)

Sample design List household frame count variable used in statistical analyses.

**For the following Sample Weight variable(s): Please see HCHS/SOL Manuscript "Sample design and cohort selection in the Hispanic Community Health Study/Study of Latinos" (Annals of Epidemiology, Volume 20, Issue 8, August 2010, Pages 629-641) for

information on how the sample weights were created. Also see the HCHS/SOL Analytic Methods document that describes the proper use of sample weights in statistical analyses.

1.8 Weight_Final_Expanded (Expanded Census 2010 Calibrated, Trimmed, Nonresponse Adjusted Weights - Year 1-3 (frozen as of 4/25/2012))

The calibrated sample weight is the calibrated, trimmed, non-response adjusted reciprocal of a participant's probability of selection into the HCHS/SOL study. The trimmed, non-response adjusted sample weights were calibrated using first the 7-category Hispanic Background distribution and then the 12-category age group/gender distribution using the 2010 Decennial Census. This variable is used for analyses purposes only.

1.9 Weight_Final_Norm_Overall (Overall Normalized Census 2010 Calibrated Trimmed NonResponse Adjusted Final Weights - Year 1-3 (frozen as of 11/30/2012))

The 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 HCHS/SOL 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. The non-response adjusted sample weights were then trimmed using the selected average percentile that minimizes model MSE 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 first the 7-category Hispanic Background distribution and then the 12-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 HCHS/SOL sample to normalize the expanded calibrated sample weights (WEIGHT_Final_Norm_Overall). This final sample weight variable is used for analyses purposes only.

1.10 Weight_Final_Norm_Center (FC-specific Normalized Census 2010 Calibrated Trimmed NonResponse Adjusted Final Weights - Year 1-3 (frozen as of 4/25/2012))

The FC-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 HCHS/SOL 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 (WEIGHT_Final_Norm_Center). 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 variable(s):

SUBJID. Subject ID

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 variable(s):

CENTER. Center (Participant's field 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 Informed Consent tracking form (ICT). If the form or the date of form completion is not present, then it is derived from the Personal Information (PIE), Anthropometry (ANT), or Sitting Blood Pressure (SBP) forms.

CLINDATE= ICTA0A or ICTB0A

or

Minimum (PIEA0A, SBPA0A, ANTA0A) if ICTA0A/ICTB0A is missing

Source variable(s):

PIEA0A. Date of completion of the Personal Information Form

ICTA0A or ICTB0A. Date of completion of the Informed Consent Form (version A or B)

SBPA0A. Date of completion of the Sitting Blood Pressure Form

ANTA0A. Date of completion of the Anthropometry Form

2.4 Consent (Informed consent status for participation in study)

This is a binary variable that determines whether or not a study participant consented to participate in the baseline examination study.

If ICTA1=1 or ICTB1=1 then consent=1;

Else consent=0;

Source variable(s):

ICTA1 or ICTB1. Agrees to participant in HCHS/SOL study (yes/no)

2.5 Full_AFU_Eligible (Core Baseline Requirements for AFU Eligibility)

A 0/1 numeric variable that defines a participant as meeting all requirements to determine if eligible for annual follow-up. The minimum subset of required forms and procedures are as follows (permanently missing forms are NOT included to determine completeness):

Questionnaires:

Informed Consent (version A or B)
Personal Information (PIE/PIS)
Personal Identifiers (IDE/IDS)
Medical History (MHE/MHS)

Procedures:

Anthropometry (ANT)
Sitting Blood Pressure (SBP)
Biospecimen Collection (BIO)

Additional criterion to be considered eligible for annual follow-up are as follows:

- Consent status to participate in Baseline Examination: Must have ICTA1=1 or ICTB1=1.
- Consent status to participate in yearly Annual Follow-Up: For those participants with an ICTA form, they must have ICTA3=1.
- Sample Blood Draws: At least one sample blood draw was collected. From the response matrix on the BIOA form, questions BIOA14A1-14A10 refer to samples not drawn for each tube (1-10). All 10 must be checked which indicates that no blood samples were drawn from participant.

Source variable(s):

ICTA1. Agrees to participate in baseline examination.

ICTA3. Agrees to yearly contact to answer questions about health/update contact information.

ICTB1. Agrees to participate in baseline examination and yearly contact and to answer questions about health/update contact information.

BIOA14A1-BIOA14A10. Blood drawing incidents – sample not drawn (tubes 1-10)

2.6 Genotyping (Full consent for genotyping)

This is a binary variable that determines whether or not a study participant fully consented to genotyping for use with investigators associated or not associated to HCHS/SOL and specialized laboratories (Note: does NOT reflect consent for use by commercial or for-profit companies that are not part of HCHS/SOL).

If ((ictb4=1 and ictb7=1 and ictb8=1) or (icta4=1 and icta6=1)) then genotyping=1.
Otherwise genotyping=0.

Source variable(s):

ICTA1 or ICTB1. Agrees to participate in HCHS/SOL.

ICTA4 or ICTB4. Agrees to provide a blood sample for DNA and RNA that can be extracted, stored and used

ICTA6. Agrees to share “de-identified” blood, genetic material (DNA/RNA) and other information with HCHS/SOL researchers and their collaborators

ICTB7. Agrees to share non-genetic data, information, and samples available to investigators not associated to HCHS/SOL and specialized laboratories

ICTB8. Agrees to share genetic data, information, and samples available to investigators not associated to HCHS/SOL and specialized laboratories

2.7 DNA_Available (DNA Available (1=Yes,0=No))

This is a binary variable that determines whether or not a study participant had blood draws successfully conducted without incidents or problems for tubes 4 and 5 to make DNA available to investigators associated or not associated to HCHS/SOL and specialized laboratories (Note: does NOT reflect consent for use on DNA samples, i.e. genotyping).

```
((bioa14a4^=1 & bioa14b4^=1 & bioa14c4^=1 & bioa14d4^=1 & bioa14e4^=1 &
  bioa14f4^=1 &
  bioa22a4^=1 & bioa22b4^=1 & bioa22c4^=1 & bioa22d4^=1 & bioa22e4^=1 ) |
(bioa14a5^=1 & bioa14b5^=1 & bioa14c5^=1 & bioa14d5^=1 & bioa14e5^=1 &
  bioa14f5^=1 &
  bioa22a5^=1 & bioa22b5^=1 & bioa22c5^=1 & bioa22d5^=1 & bioa22e5^=1 ))
```

then let DNA_Available=1

Otherwise DNA_Available=0

Source variable(s):

BIOA14A4. Tube 4 – Sample not drawn

BIOA14B4. Tube 4 – Partial sample drawn

BIOA14C4. Tube 4 – Tourniquet reapplied

BIOA14D4. Tube 4 – Fist clenching

BIOA14E4. Tube 4 – Needle movement

BIOA14F4. Tube 4 – Participant reclining

BIOA14A5. Tube 5 – Sample not drawn

BIOA14B5. Tube 5 – Partial sample drawn

BIOA14C5. Tube 5 – Tourniquet reapplied

BIOA14D5. Tube 5 – Fist clenching

BIOA14E5. Tube 5 – Needle movement

BIOA14F5. Tube 5 – Participant reclining

BIOA22A4. Tube 4 – Broken Tube

BIOA22B4. Tube 4 – Sample re-centrifuged

BIOA22C4. Tube 4 – Clotted

BIOA22D4. Tube 4 – Hemolyzed

BIOA22E4. Tube 4 – Lipemic

BIOA22A5. Tube 4 – Broken Tube

BIOA22B5. Tube 4 – Sample re-centrifuged

BIOA22C5. Tube 4 – Clotted

BIOA22D5. Tube 4 – Hemolyzed
BIOA22E5. Tube 4 – Lipemic

2.8 NP_Permit (Permission for non-profit use of data)

This is a binary variable that determines whether or not a study participant gave permission for non-profit use of study data.

ICTA7 = 1 or ICTB9 = 1 → 1
otherwise 0

Response Format: 0 = use not permitted
1 = permitted

Source variable(s):

ICTA7. Agrees to share de-identified blood, genetic material and other information with nonprofit private companies

ICTB9. Agrees to share genetic and non-genetic data with commercial or for-profit companies

2.9 External_Permit (Permission for investigators outside HCHS/SOL)

This is a binary variable that determines whether or not a study participant gave permission for use for investigators outside HCHS/SOL.

ICTA6 = 1 or ICTB7 = 1 → 1
otherwise 0

Response Format: 0 = use not permitted
1 = permitted

Source variable(s):

ICTA6. Agrees to share de-identified blood, genetic material and other information with HCHS/SOL researchers and other individuals

ICTB7. Agrees to share non-genetic data to investigators not associated to HCHS/SOL and specialized laboratories

2.10 Commercial_Permit (Permission for for-profit use of data)

This is a binary variable that determines whether or not a study participant gave permission for use for for-profit use of study data. Note that for-profit permission was not collected in ICT version A.

ICTB9 = 1 → 1
otherwise 0

Response Format: 0 = use not permitted
1 = permitted

Source variable(s):

ICTB9. Agrees to share genetic and non-genetic data with commercial or for-profit companies

2.11 Any_Permit (Any permission for public use data)

This is a binary variable that combines the 3 permission variables to determine whether or not a study participant gave any permission for public use of data.

NP_Permit=1 or External_Permit=1 or Commercial_Permit=1 → 1
otherwise 0

Response Format: 0 = use not permitted
1 = permitted

Source variable(s):

NP_Permit. Permission for non-profit use of data

External_Permit. Permission for investigators outside HCHS/SOL

Commercial_Permit. Permission for for-profit use of data

3. SOCIO-DEMOGRAPHIC

3.1 Age (Age in years at the time of participant's clinic visit)

This is the age of the 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 (PIEA2– PIEA0A)/365.25

When missing use age collected as recruitment:

ELEA2 and if this missing, then use

HSRA5A4-HSRA5N4 or HSRB4A4-HSRB4J4 based on linkage in ELEA of participant to roster row of household screening form

Source variable(s):

PIEA0A. Date of completion of the Personal Information Form

PIEA2. Date of birth

ELEA2 – Age of person screened

HSRA5A4-5N4 – Age of person rostered from household (rows A-N)

HSRB4A4-4J4 – Age of person rostered from household (rows A-J)

3.2 Agegroup_C2 (2-level grouped age of participant)

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

AGE	AGEGROUP_C2
18 – 44	1
45+	2

Response Format: 1=Ages 18-44

2=Ages 45+

Source variable(s):

AGE. age of participant at the time of their clinic visit

3.3 Agegroup_C3 (3-level grouped age of participant)

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

AGE	AGEGROUP_C2
18 – 44	1
45 – 64	2
65+	3

Response Format: 1=Ages 18-44
2=Ages 45-64
3=Ages 65+

Source variable(s):

AGE. age of participant at the time of their clinic visit

3.4 Agegroup_C6 (6-level grouped age of participant)

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

AGE	AGEGROUP_C6
18 – 24	1
25 – 34	2
35 – 44	3
45 – 54	4
55 – 64	5
65 +	6

Response Format: 1=Ages 18-24
2=Ages 25-34
3=Ages 35-44
4=Ages 45-54
5=Ages 55-64
6=Ages 65+

Source variable(s):

AGE. age of participant at the time of their clinic visit

3.5 Agegroup_C6_NHANES (6-level grouped age – NHANES standardization)

This is the categorical (grouped) age of the participant. It is determined from the derived variable AGE for use with analyses requiring NHANES age standardization (Note: Does not include participants 75+ years old at baseline).

AGE	Agegroup_C6_NHANES
18 – 29	1
30 – 39	2
40 – 49	3
50 – 59	4
60 – 69	5
71 – 74	6

Response Format: 1=Ages 18-29
2=Ages 30-39
3=Ages 40-49
4=Ages 50-59
5=Ages 60-69
6=Ages 70-74

Source variable(s):

AGE. age of participant at the time of their clinic visit

3.6 Gender (Gender (F=Female, M=Male))

This is a categorical character variable which describes the participant’s gender, female (F) or male (M).

PIEA1	Gender
1	M
2	F

When missing use gender collected at recruitment:

HSRA5A3-HSRA5N3 or HSRB4A3-HSRB4J3 based on linkage in ELEA of participant to roster row of household screening form

Response Format: M=Male
F=Female

Source variable(s):

PIEA1. Gender

HSRA5A3-5N3 – Gender of person rostered from household (rows A-N)

HSRB4A3-4J3 – Gender of person rostered from household (rows A-J)

3.7 Gendernum (Gender – Numeric (0=Female, 1=Male))

This is a 0/1 variable which describes the participant’s gender, female (0) or male (1).

Gender	Gendernum
F	0
M	1

Response Format: 0=Female
1=Male

Source variable(s):

Gender. (Gender (F=Female, M=Male))

3.8 Education_C2 (2-level group education level)

The educational status derived variable groups the attainment of a high school diploma or an equivalent degree, or not having that degree. Update 12/2013: categorized missing values based on PIEA21B-PIEA21G, PIEA19B, PIEA21I, and PIEA17.

Ordered categorization for Response Format

1. Highest level of elementary, middle school or high school (PIEA19A=1 or 2 or 3)
AND no high school diploma or GED (PIEA21A=0 or missing) → 1

2. [Highest level of elementary, middle school or high school (PIEA19A=1 or 2 or 3) AND high school diploma or GED (PIEA21A=1)] OR Highest level of trade school/vocational school or university/college (PIEA19A=4 or 5) → 2
3. 0 years of education (PIEA17=0) → 1
4. Trade school certificate, associate degree, bachelor's degree, etc. (Any of PIEA21B through PIEA21G=1) → 2
5. Text response for other level indicating No high school diploma or GED (PIEA19B) → 1
6. Text response for other level indicating At least High school diploma or GED (PIEA19B) → 2
7. Text response for other degree indicating At least High school diploma or GED (PIEA21I) → 2
8. 1-10 years of education ($1 \leq \text{PIEA17} \leq 10$) → 1
9. 11+ years of education ($\text{PIEA17} \geq 11$) → 2
10. Otherwise missing

Response format: 1 = No high school diploma or GED
2 = High school diploma/GED or greater

Source variable(s):

- PIEA17. How many years of schooling in total have you completed?
- PIEA19A. Highest (or closest) grade/level of education achieved (6 levels)
- PIEA19B. If other, please specify:
- PIEA21A. High school diploma or equivalent (yes/no)
- PIEA21B. Trade school/vocational school certificate
- PIEA21C. Associate degree (i.e. AA, AS)
- PIEA21D. Bachelor's degree (i.e. BA, AB, BS)
- PIEA21E. Master's degree (i.e. MA, MS, MEd, MSW, MBA)
- PIEA21F. Professional degree (i.e. MD, DDS, DVM, LLB, JD)
- PIEA21G. Doctorate degree (i.e. PhD, EdD)
- PIEA21I. If other, please specify:

3.9 Education_C3 (3-level group education level)

The educational status derived variable groups the attainment of a high school diploma or an equivalent degree, not having that high school or equivalent degree, or attainment of education beyond a high school equivalent (i.e. college or vocational). Update 12/2013: categorized missing values based on PIEA21B-PIEA21G, PIEA19B, PIEA21I, and PIEA17.

Ordered categorization

1. Highest level of elementary, middle school or high school (PIEA19A=1 or 2 or 3) AND no high school diploma or GED (PIEA21A=0 or missing) → 1

2. Highest level of elementary, middle school or high school (PIEA19A=1 or 2 or 3) AND high school diploma or GED (PIEA21A=1) → 2
3. Highest level of trade school/vocational school or university/college (PIEA19A=4 or 5) → 3
4. 0 years of education (PIEA17=0) → 1
5. Trade school certificate, associate degree, bachelor's degree, etc. (Any of PIEA21B through PIEA21G=1) → 3
6. Text response for other level indicating No high school diploma or GED (PIEA19B) → 1
7. Text response for other level indicating At most a High school diploma or GED (PIEA19B) → 2
8. Text response for other level indicating Greater than high school (or GED) education (PIEA19B) → 3
9. Text response for other degree indicating At most a High school diploma or GED (PIEA21I) → 2
10. Text response for other degree indicating Greater than high school (or GED) education (PIEA21I) → 3
11. 1-10 years of education ($1 \leq \text{PIEA17} \leq 10$) → 1
12. 11-12 years of education ($11 \leq \text{PIEA17} \leq 12$) → 2
13. 13+ years of education ($\text{PIEA17} \geq 13$) → 3
14. Otherwise missing

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

Source variable(s):

- PIEA17. How many years of schooling in total have you completed?
- PIEA19A. Highest (or closest) grade/level of education achieved (6 levels)
- PIEA19B. If other, please specify:
- PIEA21A. High school diploma or equivalent (yes/no)
- PIEA21B. Trade school/vocational school certificate
- PIEA21C. Associate degree (i.e. AA, AS)
- PIEA21D. Bachelor's degree (i.e. BA, AB, BS)
- PIEA21E. Master's degree (i.e. MA, MS, MEd, MSW, MBA)
- PIEA21F. Professional degree (i.e. MD, DDS, DVM, LLB, JD)
- PIEA21G. Doctorate degree (i.e. PhD, EdD)
- PIEA21I. If other, please specify:

3.10 Education_C4 (4-level group education level)

Based on Education_C3, Education_C4 was created to differentiate between some colleges vs MORE than college.

If Education_C3=3 and PIEA19a=5 then set Education_C4=4.

Otherwise Education_C4=Education_C3

Response format: 1 = No high school diploma or GED
2 = At most a High school diploma or GED
3 = High school (or GED) education
4 = University/college education

Source variable(s):

Education_C3. 3-level group education level.

PIEA19A. Highest (or closest) grade/level of education achieved (6 levels)

3.11 Income (Grouped yearly household income)

This derived variable groups each subject to one income category. The income is set to missing if both ECEA3 and ECEA4 are missing. Income is set to ECEA3 if non-missing. Otherwise income is set to ECEA4.

Income	ECEA3	ECEA4
1	1	.
2	2	.
3	3	.
4	4	.
5	5	.
6	.	1
7	.	2
8	.	3
9	.	4
10	.	5
.	.	.

Response format: 1 = Less than \$10,000
2 = \$10,001-\$15,000
3 = \$15,001-\$20,000
4 = \$20,001-\$25,000
5 = \$25,001-\$29,999
6 = \$30,000-\$40,000
7 = \$40,001-\$50,000
8 = \$50,001-\$75,000
9 = \$75,001-\$100,000
10 = More than \$100,000

Source variable(s):

ECEA3. Low income range

ECEA4. High income range

3.12 Income_C3 (3-level grouped yearly household income)

This derived variable groups each subject to one income category based exclusively on ECEA2. ECEA2 was answered by some participants who refused to answer the follow up questions ECEA3 and ECEA4. Participants with missing ECEA2 grouped together.

ECEA2=1 → 1
ECEA2=2 → 2
ECEA2 is missing or zero → 3

Response format: 1=Less than \$30,000
2=\$30,000 or more
3=Missing

Source variable(s):

ECEA2. Your household income for the year

3.13 Income_C5 (5-level grouped yearly household income)

This derived variable groups each subject to one income category. The income is set to missing if both ECEA3 and ECEA4 are missing. Income_C5 is set to ECEA3 if non-missing. Otherwise income is set to ECEA4.

```
if ecea3 in(1) then INCOME_C5=1;  
else if ecea3 in(2,3) then INCOME_C5=2;  
else if ecea3 in(4,5) | ecea4=1 then INCOME_C5=3;  
else if ecea4 in(2,3) then INCOME_C5=4;  
else if ecea4 in(4,5) then INCOME_C5=5;  
else INCOME_C5=.;
```

Response format: 1 = Less than \$10,000
2 = \$10,001-\$20,000
3 = \$20,001-\$40,000
4 = \$40,001-\$75,000
5 = More than \$75,000

Source variable(s):

ECEA3. Low income range
ECEA4. High income range

3.14 POVPCT (Household income, percent of poverty)

This is a numeric variable that calculates the ratio of household income in US dollars by poverty threshold, as follows:

Intermediate Variable for POVPCT:

```
if income=1 then INCOME_DOLLARS=5000;  
else if income=2 then INCOME_DOLLARS=12500;  
else if income=3 then INCOME_DOLLARS=17500;  
else if income=4 then INCOME_DOLLARS=22500;
```


else if income=5 then INCOME_DOLLARS=27500;
else if income=6 then INCOME_DOLLARS=35000;
else if income=7 then INCOME_DOLLARS=45000;
else if income=8 then INCOME_DOLLARS=62500;
else if income=9 then INCOME_DOLLARS=87500;
else if income=10 then INCOME_DOLLARS=105000;

Intermediate Variable for POVPCT:

clindate_year=year(clindate);
if ecea5=1 & age<65 & clindate_year=2008 then POVERTY_THRESHOLD=11201;
else if ecea5=1 & age<65 & clindate_year=2009 then POVERTY_THRESHOLD=11161;
else if ecea5=1 & age<65 & clindate_year=2010 then POVERTY_THRESHOLD=11344;
else if ecea5=1 & age<65 & clindate_year=2011 then POVERTY_THRESHOLD=11702;

else if ecea5=1 & age>=65 & clindate_year=2008 then POVERTY_THRESHOLD=10326;
else if ecea5=1 & age>=65 & clindate_year=2009 then POVERTY_THRESHOLD=10289;
else if ecea5=1 & age>=65 & clindate_year=2010 then POVERTY_THRESHOLD=10458;
else if ecea5=1 & age>=65 & clindate_year=2011 then POVERTY_THRESHOLD=10788;

else if ecea5=2 & age<65 & clindate_year=2008 then POVERTY_THRESHOLD=14489;
else if ecea5=2 & age<65 & clindate_year=2009 then POVERTY_THRESHOLD=14439;
else if ecea5=2 & age<65 & clindate_year=2010 then POVERTY_THRESHOLD=14676;
else if ecea5=2 & age<65 & clindate_year=2011 then POVERTY_THRESHOLD=15139;

else if ecea5=2 & age>=65 & clindate_year=2008 then POVERTY_THRESHOLD=13030;
else if ecea5=2 & age>=65 & clindate_year=2009 then POVERTY_THRESHOLD=12982;
else if ecea5=2 & age>=65 & clindate_year=2010 then POVERTY_THRESHOLD=13194;
else if ecea5=2 & age>=65 & clindate_year=2011 then POVERTY_THRESHOLD=13609;

else if ecea5=3 & clindate_year=2008 then POVERTY_THRESHOLD=17163;
else if ecea5=3 & clindate_year=2009 then POVERTY_THRESHOLD=17098;
else if ecea5=3 & clindate_year=2010 then POVERTY_THRESHOLD=17374;
else if ecea5=3 & clindate_year=2011 then POVERTY_THRESHOLD=17916;

else if ecea5=4 & clindate_year=2008 then POVERTY_THRESHOLD=22025;
else if ecea5=4 & clindate_year=2009 then POVERTY_THRESHOLD=21954;
else if ecea5=4 & clindate_year=2010 then POVERTY_THRESHOLD=22314;
else if ecea5=4 & clindate_year=2011 then POVERTY_THRESHOLD=23021;

else if ecea5=5 & clindate_year=2008 then POVERTY_THRESHOLD=26049;
else if ecea5=5 & clindate_year=2009 then POVERTY_THRESHOLD=25991;
else if ecea5=5 & clindate_year=2010 then POVERTY_THRESHOLD=26439;
else if ecea5=5 & clindate_year=2011 then POVERTY_THRESHOLD=27251;

else if ecea5=6 & clindate_year=2008 then POVERTY_THRESHOLD=29456;
else if ecea5=6 & clindate_year=2009 then POVERTY_THRESHOLD=29405;
else if ecea5=6 & clindate_year=2010 then POVERTY_THRESHOLD=29897;
else if ecea5=6 & clindate_year=2011 then POVERTY_THRESHOLD=30847;

else if ecea5=7 & clindate_year=2008 then POVERTY_THRESHOLD=33529;
else if ecea5=7 & clindate_year=2009 then POVERTY_THRESHOLD=33372;
else if ecea5=7 & clindate_year=2010 then POVERTY_THRESHOLD=34009;
else if ecea5=7 & clindate_year=2011 then POVERTY_THRESHOLD=35085;

else if ecea5=8 & clindate_year=2008 then POVERTY_THRESHOLD=37220;
else if ecea5=8 & clindate_year=2009 then POVERTY_THRESHOLD=37252;
else if ecea5=8 & clindate_year=2010 then POVERTY_THRESHOLD=37934;
else if ecea5=8 & clindate_year=2011 then POVERTY_THRESHOLD=39064;

else if ecea5>=9 & clindate_year=2008 then POVERTY_THRESHOLD=44346;
else if ecea5>=9 & clindate_year=2009 then POVERTY_THRESHOLD=44366;
else if ecea5>=9 & clindate_year=2010 then POVERTY_THRESHOLD=45220;
else if ecea5>=9 & clindate_year=2011 then POVERTY_THRESHOLD=46572;

Define POV PCT:

POVPCT=round((income_dollars/poverty_threshold)*100);

Source variable(s):

Income. Grouped yearly household income.
Clindate. HCHS/SOL baseline visit date.
ECEA5. Household family member number.

3.15 POV PCT_C4 (3-level grouped Household income, percent of poverty)

This derived variable groups each subject's Household income, percent of poverty as follow.

if .Z<povpct<=100 then POV PCT_C4=1;
else if 101<=povpct<=200 then POV PCT_C4=2;
else if 201<=povpct<=300 then POV PCT_C4=3;
else if povpct>300 then POV PCT_C4=4;

Response format: 1=Less than 100
2=101 to 200
3=201 to 300
4=Greater than 300

Source variable(s):

POVPCT. Household income, percent of poverty

3.16 Marital_Status (4-level grouped marital status)

This derived variable groups each subject's marital status.

if piea3=1 then MARITAL_STATUS=1;
else if piea3 in (2,6) then MARITAL_STATUS=2;
else if piea3 in (3,4,5) then MARITAL_STATUS=3;
else MARITAL_STATUS=.;

Response format: 1=Single
2=Married or living with a partner
3=Separated, divorced, or widow(er)

Source variable(s):

PIEA3. Marital Status

3.17 Employed (4-level grouped employment status)

This variable groups the employment status of the participant. This variable includes retirees who are or are not employed. Retired participants missing employment are categorized as not currently employed, while unemployed participants missing data on retirement are categorized as not retired.

```
if (oce1=1 & (oce5=3 | oce5≤.Z)) then EMPLOYED=1;
else if (oce1^=1 & oce5=3) then EMPLOYED=2;
else if (oce5=2) then EMPLOYED=3;
else if (oce5=1) then EMPLOYED=4;
else EMPLOYED=.;
```

Response Format: 1=Retired and not currently employed (or missing on employment).
2=Not retired (or missing on retirement) and not currently employed
3=Employed part-time (≤35 hours/week)
4=Employed full-time (>35 hours/week)

Source variable(s):

OCEA1. Retired (yes/no)

OCEA5. Current employment status

3.18 N_HC (Health Insurance Coverage - Current)

This variable groups the health insurance coverage status of the participant. The definition also checks whether or not any reason was giving for stopping coverage.

```
hi_ind=sum(hcea10b,hcea10c,hcea10d,hcea10e,hcea10f,hcea10g,hcea10h,hcea10i,hcea10j);
```

```
stop_ind=sum(hcea12a,hcea12b,hcea12c,hcea12d,hcea12e,hcea12f,hcea12g,hcea12h,hcea12i,hcea12j,hcea12k);
```

```
if hcea10a≤.Z and hi_ind>0 then N_HC=1;
else if hcea10a≤.Z and stop_ind>0 then N_HC=0;
else if hceb10a=1 then N_HC=1;
else if hceb10a=0 then N_HC=0;
```

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

Source variable(s):

HCEA10a/HCEB10a. Do you have health insurance or other health care coverage?

HCEA10b/HCEB10b-HCEA10j/HCEB10j. What type of health insurance or health care coverage do you have?

HCEA12a/HCEB12a-HCEA12k/HCEB12k. Reasons you stopped being covered by health insurance.

3.19 Occupation_Curr (Grouped current occupation for employed participants)

For participants who are currently employed (part or full time), this is the type of occupation the participant had at the time of the baseline examination.

oce16a=7 → 1
oce16a=11 → 2
oce16a=6 → 3
1≤oce16a≤4 → 4;
other nonmissing values → 5

Response format: 1=Non-skilled worker
2=Service worker
3=Skilled worker
4=Professional/technical, administrative/executive, or office staff
5=Other occupation

Source variable(s):

OCEA16a. Type of job currently work majority of hours per week.

3.20 Occupation_Long (Grouped longest held occupation)

This is the type of occupation the participant has held the longest.

oce28a=7 → 1
oce28a=11 → 2
oce28a=6 → 3
1≤oce28a≤4 → 4;
other nonmissing values → 5

Response format: 1=Non-skilled worker
2=Service worker
3=Skilled worker
4=Professional/technical, administrative/executive, or office staff
5=Other occupation

Source variable(s):

OCEA28a. Type of job held the longest.

4. ACCULTURATION

4.1 Bkgrd1 (Re-classification of Hispanic/Latino Background)

The self-identification with a Hispanic/Latino cultural group allowed for selection of an 'other' designation. This variable groups the text responses of 7=Other (using PIEA5B) for variations of 'Dominican' so that undercounts do not occur for that group. In addition, several 7=Other responses are re-categorized into one of the 7 groups as defined below.

Response format:

- 0 = Dominican
- 1 = Central American
- 2 = Cuban
- 3 = Mexican
- 4 = Puerto Rican
- 5 = South American
- 6 = More than one heritage
- 7 = Other

Where 7=Other, the responses to PIEA5B below have BKGRD1 redefined as follows (as of 9/30/2013):

BKGRD1	PIEA5B text response
0 - Dominican heritage background	'DOIMINICAN', 'DOMENICANA', 'DOMENICANO', 'DOMICANO', 'DOMINCAN', 'DOMINICAN', 'DOMINICAN DESCENT', 'DOMINICAN REPUBLIC', 'DOMINICANA', 'DOMINICANO', 'REPUBLICA DOMINICA', 'SANTO DOMINGO' 'HISPANIC/DOMINICAN', 'DOMINICAN/GERMAN D', 'REPUBLIC SANTO DOM'
1 - Central American heritage background	'HUNDURAS', 'HONDURENIA', 'SALVADORENA', 'GUATEMALAN', 'JUST EL SALVADOR', 'CENTROAMERICANA', 'NICARAGUENCE'
2 - Cuban heritage background	'2' (Reason: '2' = Cuban heritage in PIEA5a options)
3 - Mexican heritage background	'CHICANA', 'CHICANO', 'MEXICAN', 'MEXICAN AMERICAN', 'MEXICAN-AMERICAN', 'MEXICO AMERICANA', 'ASIATICO/MEXICANO'

4 - Puerto-Rican heritage background	'NEW YORK RICAN', 'USA PUERTORICAN', 'PARCIAL PR'
5 - South American heritage background	'BRASIL', 'COLOMBIANA', 'ECUATORIANO', 'PERUANA', 'SPAIN/BRAZIL', 'JUST ECUADORIAN', 'ECUADOR (SEE NOTE)', 'ARGENTINA'
6 - More than one heritage	'CUBAN/COLOMBIAN', 'MEXICAN/PR', 'MIXED RACE', 'MEXICAN/PUERTO RIC', 'EQUADORIAN/P.RICAN', 'PR AND DOMINICAN', 'PR & DOMINICAN', 'PR AND ECUADORIAN', 'BOTH 3 AND 4'

Source variable(s):

PIEA5A. Hispanic/Latino background

PIEA5B. Text response for other Hispanic/Latino background

4.2 Bkgrd1_C7 (7-level re-classification of Hispanic/Latino Background)

Variable as defined for Bkgrd1. Also, the Other and More than one heritage categories are combined into one group.

Response format: 0 = Dominican
1 = Central American
2 = Cuban
3 = Mexican
4 = Puerto Rican
5 = South American
6 = More than one/Other heritage

Source variable(s):

PIEA5A. Hispanic/Latino background

PIEA5B. Text response for other Hispanic/Latino background

4.3 Bkgrd1_C6 (6-level re-classification of Hispanic/Latino Background)

Variable as defined for Bkgrd1. Also, Central and South American categories are combined into one group and the Other and More than one heritage categories are combined into one group.

Response format: 0 = Dominican
1 = Central or South American
2 = Cuban
3 = Mexican
4 = Puerto Rican
5 = More than one/Other heritage

Source variable(s):

PIEA5A. Hispanic/Latino background

PIEA5B. Text response for other Hispanic/Latino background

4.4 Site_Bkgrd (Cross-Classification of Center and Hispanic/Latino Background)

The self-identification with a Hispanic/Latino cultural group crossed by field center and grouped into the following categories:

Response format: 1 = Dominicans - Bronx
2 = Central American - Bronx
3 = Central American - Chicago
4 = Central American - Miami
5 = Cubans - Miami
6 = Mexicans - Bronx
7 = Mexicans - Chicago
8 = Mexicans - San Diego
9 = Puerto Ricans - Bronx
10 = Puerto Ricans - Chicago
11 = South American - Bronx
12 = South American - Chicago
13 = South American - Miami
14 = Other - Bronx
15 = Other - Chicago
16 = Other - Miami
17 = Other - San Diego

Source variable(s):

CENTER. Field Center

BKGRD1_C7. Hispanic/Latino background

CENTER (Center)	BKGRD1_C7(7-level re-classification of Hispanic/Latino Background)						
	Dominican	Central American	Cuban	Mexican	Puerto Rican	South American	Mixed/ Other
Bronx	1	2	14	6	9	11	14
Chicago	15	3	15	7	10	12	15
Miami	16	4	5	16	16	13	16
San Diego	17	17	17	8	17	17	17

13 Grey shaded boxes indicate adequate sample size ($n \geq 100$) for comparisons between backgrounds within sites or vice versa. All backgrounds with < 100 within each site are pooled with the Mixed/other category for that site.

4.5 Site_Bkgrd_45Plus (Cross-Classification of Center and Hispanic/Latino Background, 45+ Years Only)

The self-identification with a Hispanic/Latino cultural group, based on a sub-cohort of participants 45+ years old, crossed by field center and grouped into the categories below. Participants younger than 45 are classified as missing.

Response format:

- 1 = Dominicans - Bronx
- 2 = Central American - Bronx
- 3 = Central American - Chicago
- 4 = Central American - Miami
- 5 = Cubans - Miami
- 6 = Mexicans - Chicago
- 7 = Mexicans - San Diego
- 8 = Puerto Ricans - Bronx
- 9 = Puerto Ricans - Chicago
- 10 = South American - Bronx
- 11 = South American - Chicago
- 12 = South American - Miami
- 13 = Other – Bronx (including missing on BKGRD1_C7)
- 14 = Other – Chicago (including missing on BKGRD1_C7)
- 15 = Other – Miami (including missing on BKGRD1_C7)
- 16 = Other – San Diego (including missing on BKGRD1_C7)

Source variable(s):

CENTER. Field Center

BKGRD1_C7. Hispanic/Latino background

AGE. Age in years at the time of participant's clinic visit

CENTER (Center)	BKGRD1_C7(7-level re-classification of Hispanic/Latino Background)						
	Dominican	Central American	Cuban	Mexican	Puerto Rican	South American	Mixed/ Other
Bronx	1	2	13	13	8	10	13
Chicago	14	3	14	6	9	11	14
Miami	15	4	5	15	15	12	15
San Diego	16	16	16	7	16	16	16

12 Grey shaded boxes indicate adequate sample size ($n \geq 100$) for comparisons between backgrounds within sites or vice versa. All backgrounds with < 100 within each site are pooled with the Mixed/other category for that site.

4.6 Race (Race Group (self-report))

The self-identification with a race group. Refusals (.Q) have been coded to a valid category (9) and unknown or missing values have been coded to another valid value (7).

If PIEA6 = .Q then RACE = 9
 Else if PIEA6 = (. , 7) then RACE=7
 Else then RACE=PIEA6.

Response format:

- 1 = "Am. Indian/AK native"
- 2 = "Asian"
- 3 = "HW native/Pac. Islander"
- 4 = "Black"
- 5 = "White"
- 6 = "Multi-racial"
- 7 = "Unknown/NR"
- 9 = "Refused"

Source variable(s):

PIEA6. Participant's self-reported race

4.7 US_Born (Grouped place of birth – 50 US States only)

This variable groups the place of birth of the participant to the United States (50 states only) or other place of birth (defines refused responses as missing).

if piea4='63' then US_BORN=1;
 else if piea4^='63' & piea4 not in(' ', 'Q', 'q') then US_BORN=0;
 else US_BORN=.;

Response format: 0=Not born in 50 US States.
 1=Born in 50 US States Only

Source variable(s):

PIEA4. Born in which Country/territory.

4.8 US_Native (Grouped place of birth – US states and territories)

This variable groups the place of birth of the participant to the United States (50 states and US territories which includes Puerto-Rico, Guam, and Virgin Islands) or other place of birth (defines refused responses as missing).

```
if piea4 in(' ','Q','q') then US_NATIVE=.;  
else if piea4 in('29','57','63','66') then US_NATIVE=1;  
else US_NATIVE=0;
```

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

Source variable(s):

PIEA4. Born in which Country/territory.

4.9 Lang_Pref (Language Preference for Baseline Examination)

This variable determines which language was preferred to be used for the baseline examination. Three core forms are used: Personal Information (PIEA), Personal Identifiers (IDEA), and Medical History (MHEA). A majority of those used in one language determines which language was preferred.

Response format: 1=Spanish
2=English

Source variable(s):

Form Version of PIEA/PISA
Form Version of IDEA/IDSA
Form Version of MHEA/MHSA

4.10 IMGEN_C2 (2-level Immigrant Generation)

This variable determines the participants immigrant generation status based on a combination of the country of origin of the participant, maternal and paternal parents, and/or the maternal and paternal grandparents as follows:

First generation is defined as foreign-born with foreign-born parents (i.e., two foreign-born parents, or one foreign-born parent and one parent missing; same as IMGEN_C4).

The second generation and beyond definition:

If US_BORN in (0, 1) **with both** PIEA11 and PIEA14 are missing, then set IMGEN_C2 as missing.

Else if US_BORN=1 then classified as 2nd + generation.

Else if US_BORN =0 **with at least one US born parent** then classified as 2nd + generation

Else → set IMGEN_C2 as missing.

NOTE: Persons who are born in a US territory are treated as foreign-born or first generation immigrants. This better reflects their migration and acculturation patterns when compared to those born on the US mainland (i.e. US 50).

5/2014 update: recoded 9 to missing

Response format: 1=1st generation
2=2nd generation or higher

Source variable(s):

US_BORN. Grouped place of birth – 50 US states only

PIEA11. Mother’s country of origin

PIEA14. Father’s country of origin

4.11 IMGEN_C4 (4-level Immigrant Generation)

This variable determines the participants immigrant generation status based on a combination of the country of origin of the participant, maternal and paternal parents, and/or the maternal and paternal grandparents as follows:

First generation is defined as foreign-born with foreign-born parents (same as IMGEN_C2).

Second generation is defined as US born with **at least one** foreign-born parent. In addition, foreign-born adults with at least one US born parent are classified as second generation to reflect their mixed cultural backgrounds.

The third and fourth generation is defined as US born respondent and both US born mother and father with the following grandparents algorithm shown in chart below:

Grandparent 1	fb	fb	fb	fb	fb	fb	fb	us b	us b	us b
Grandparent 2	fb	fb	fb	fb	.	fb	fb	fb	fb	.	.	.	us b	us b	us b
Grandparent 3	fb	fb	fb	.	.	usb	us b	us b	us b	us b	.	.	us b	us b	us b
Grandparent 4	fb	usb	us b	.	us b	us b	us b	.	.	fb	us b
Generation (IMGEN_C4)	3	3	3	3	3	3	4	4	4	4	4	.	4	4	4

usb= us born, fb= foreign born as defined in IMGEN_C2

Response format: 1=1st generation
2=2nd generation
3=3rd generation
4=4th generation or higher

Source variable(s):

US_BORN. Grouped place of birth – 50 US states only

PIEA11. Mother's country of origin

PIEA12. Maternal grandmother's country of origin

PIEA13. Maternal grandfather's country of origin

PIEA14. Father's country of origin

PIEA15. Paternal grandmother's country of origin

PIEA16. Paternal grandfather's country of origin

4.12 YRSUS (Years lived in the US)

This is a numeric variable that defined years lived in the US as follows:

1. If participant was NOT born in the US, then the years lived in the US is equal to the response of PIEA7.
2. If participant was born in the US, then the years lived in the US is equal to the participant's age (calculated from participant's date of birth).

Source variable(s):

Age. Age of participant in years.

PIEA4. Participant's country of origin

PIEA7. Years lived in mainland US.

4.13 YRSUS_C2 (2-level grouped years lived in the US (50 States))

This is a 2-level grouped numeric variable that defined less than 10 years lived in the US versus 10 or more years lived in the US.

Response format: 1=Less than 10 years
2=10 years or more

Source variable(s):

YRSUS. Years lived in the US.

4.14 YRSUS_C3 (3-level grouped years lived in the US (50 States))

This is a 3-level grouped numeric variable that defined less than 10 years lived in the US versus 10, more years lived in the US or US born.

Response format: 1=Less than 10 years
2=10 years or more
3= US born

Source variable(s):

YRSUS. Years lived in the US.

US_BORN. Born in US.

4.15 SASH_ALL (Short acculturation scale for Hispanic)

5/2014 update: the variable SASH_ALL was removed from the HCHS participant derived variable file in June 2014 based on recommendations adopted by the steering committee from the HCHS sociocultural working group. SASH_ALL failed to demonstrate adequate psychometric properties in the HCHS cohort data and has been eliminated from the database. Use the two subscales from the SASH instead, SASH_LANG and SASH_SOC.

4.16 SASH_LANG (Short acculturation scale for Hispanics – language subscale)

This is a numeric variable. A 5-point likert-type scale is assigned to each question with 1 = Only Spanish and 5= Only English. The average of the 6 questions indicates the degree of language acculturation. If more than 1 item is missing, SASH_LANG is set to missing.

SASH_LANG = average (of scea1-6).

5/2014 update: Added item SCEA6, set to missing if more than 1 item is missing

Source variable(s):

SCEA1. In general, what language(s) do you read and speak?

SCEA2. What was the language(s) you used as a child?

SCEA3. What language(s) do you usually speak at home?

SCEA4. In which language(s) do you usually think?

SCEA5. What language(s) do you usually speak with your friends?

SCEA6. In general, what language(s) are the movies, T.V. and radio programs you prefer to watch and listen to?

4.17 SASH_MEDIA (Short acculturation scale for Hispanics – media)

This derived variable has been dropped on the recommendation of the HCHS socio-cultural working group. Use SCEA6 directly instead of the former SASH_Media scale

4.18 SASH_SOC (Short acculturation scale for Hispanics – Social subscale)

This is a numeric variable. A 5-point likert-type scale is assigned to each question with 1 = All Hispanic/Latino and 5= All non-Hispanic/non-Latino. The average of the 4 questions indicates the degree of social acculturation. If any item is missing, SASH_SOC is set to missing.

SASH = average (of scea7-10).

5/2014 update: set to missing if any item is missing

Source variable(s):

SCEA7. Your close friends are:

SCEA8. You prefer going to social gatherings/parties at which the people are:

SCEA9. The persons you visit or who visit you are:

SCEA10. If you could choose your children's friends, you would want them to be:

4.19 ETHIS (Ethnic Identification Score)

This is a numeric variable. A 7-point likert-type scale (values of 1.0-4.0 by 0.5 scale) is assigned to each question with 1.0 = Strongly Disagree and 4.0= Strongly Agree. The average of the 2 questions indicates the degree of ethnic identity. If either item is missing, ETHIS is set to missing.

ETHIS = average (of sceA17-18).

5/2014 update: set to missing if either item is missing

Source variable(s):

SCEA17. I have a strong sense of belonging to my own ethnic group.

SCEA18. I have a lot of pride in my ethnic group.

4.20 Language_Subscore_MESA (Language Subscore - MESA)

This is a language subscore of the acculturation score similar to the one used in MESA. It is based on the language spoken at home.

This is a numeric variable with possible values from 0 to 2 in 0.5 increments or missing. 0 indicates the lowest level of acculturation, 2 – the highest.

SCEA3	Language_Subscore_MESA
1	0.0
2	0.5
3	1.0
4	1.5
5	2.0
Missing	Missing

Response Format: 0.0 = Only Spanish
0.5 = More Spanish than English
1.0 = Both equally
1.5 = More English than Spanish
2.0 = Only English

Source variable(s):

SCEA3. What language(s) do you usually speak at home?

4.21 Nativity_Subscore_MESA (Nativity Subscore - MESA)

This is a nativity subscore of the acculturation score similar to the one used in MESA. It is based on nativity and years of residence in the U.S.

This is a numeric variable. Possible values are integers from 0 to 3 or missing. 0 indicates the lowest level of acculturation, 3 – the highest.

If US_Born is missing or (US_Born=0 and YRSUS is missing) then Nativity_Subscore_MESA is missing.

Otherwise:

Nativity_Subscore_MESA is 3, if US_Born=1;

2, if US_Born=0 and 20<=YRSUS;

1, if US_Born=0 and 10<=YRSUS<20;

0, if US_Born=0 and .z<YRSUS<10.

Source variable(s):

US_Born. Grouped place of birth – 50 US States only

YRSUS. Years lived in the US

4.22 Accult_MESA (Acculturation Score - MESA)

This is a summary acculturation score based on nativity, language spoken at home, and years of residence in the U.S. (similar to the acculturation score used in MESA). It is defined as the sum of nativity subscore and language subscore.

This is a numeric variable with possible values from 0 to 5 in 0.5 increments or missing. 0 indicates the lowest level of acculturation, 5 – the highest.

If Nativity_Subscore_MESA is missing or Language_Subscore_MESA is missing then Accult_MESA is missing;

Otherwise

Accult_MESA= Nativity_Subscore_MESA + Language_Subscore_MESA.

Source variable(s):

Nativity_Subscore_MESA. Nativity Subscore – MESA

Language_Subscore_MESA. Language Subscore - MESA

4.23 Age_IMMI (Age of immigration among participants who were not born in US mainland (50 US States + DC))

This is a numeric integer valued variable that defined age of immigration as follows:

1. If participant was NOT born in the US 50 states, then the age of immigration is equal to the participants' age minus years lived in the US.
2. If participant was born in the US 50 states, then the age of immigration is equal to missing. Note that 1 person had a missing value for US_BORN and was treated as foreign born.

For participants whose calculated age of immigration is minus 1, shift the value to zero.

```
if us_born=1 then AGE_IMMI=.;  
else AGE_IMMI=(age-piea7);  
if .Z<AGE_IMMI<0 then AGE_IMMI=0;
```

Source variable(s):

Age. Age in years at the time of participant's clinic visit

PIEA7. Years lived in mainland US.

US_Born. Grouped place of birth – 50 US States/DC only

5. ANTHROPOMETRY

5.1 BMI (Body Mass Index kg/m²)

This is a numeric variable.

Missing if the variable, HEIGHT, is missing.

If HEIGHT is not missing then $BMI = ANTA4 / (HEIGHT/100)^2$

Source variable(s):

HEIGHT. Corrected height in centimeters

ANTA4. Weight in kilograms

5.2 BMIGRP_C4 (4-level grouped Body Mass Index - WHO)

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

```
if 0≤BMI<18.5 then BMIGRP_C4=1;
else if 18.5≤BMI<25 then BMIGRP_C4=2;
else if 25≤BMI<30 then BMIGRP_C4=3;
else if 30≤BMI then BMIGRP_C4=4;
else if BMI=.Z then BMIGRP_C4=.;
```

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

Source variable(s):

BMI. Body Mass Index in kg/m²

5.3 BMIGRP_C6 (6-level grouped Body Mass Index - WHO)

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

```
if 0≤BMI<18.5 then BMIGRP_C6=1;
else if 18.5≤BMI<25 then BMIGRP_C6=2;
else if 25≤BMI<30 then BMIGRP_C6=3;
else if 30≤BMI<35 then BMIGRP_C6=4;
else if 35≤BMI<40 then BMIGRP_C6=5;
else if 40≤BMI then BMIGRP_C6=6;
else if BMI=.Z then BMIGRP_C6=.;
```

Response format: 1=Underweight (BMI < 18.5)
2=Normal (18.5 ≤ BMI < 25)
3=Overweight (25 ≤ BMI < 30)
4=Obese I (30 ≤ BMI < 35)
5=Obese II (35 ≤ BMI < 40)
6=Obese III

Source variable(s):

BMI. Body Mass Index in kg/m²

5.4 Waist_hip (Waist to Hip ratio)

This is a numeric variable.

Missing if the hip girth is missing

If hip girth is not missing then WAIST_HIP = ANTA10A / ANTA10B

Source variable(s):

ANTA10A. Waist girth measured in centimeters

ANTA10B. Hip girth measured in centimeters

5.5 Height (Height corrected in cm)

This is a numeric variable. During two time periods the standing height as measured at the Bronx was incorrect due to a mis-mounted stadiometer used in measuring all participants. The derived variable was implemented based on a recommendation from the Quality Control subcommittee.

Missing if the variable, ANTA2 was originally missing.

For participants from the Bronx,

During the time period from the start of the study until October 15, 2008:

height= ANTA2+3.2 cm

From October 16, 2008 through February 1, 2009:

height= ANTA2+2.9 cm

Otherwise, including participants from other field centers, height= ANTA2

Source variable(s):

ANTA0a date of the anthropometry measurement

ANTA2. Height in centimeters

CENTER. Participant's field center

6. RESPIRATORY

6.1 Asthma_Ever (Ever had asthma)

This is a 0/1/9 numeric variable. From the self-report on the respiratory history form, if the participant says that they ever had asthma, then:

If RSEA32 is not missing, then ASTHMA_EVER= RSEA32.

Response format: 0 = No
1 = Yes
9 = Don't Know

Source variable(s):

RSEA32. Ever had asthma

6.2 Asthma_Ever_MD (Ever had asthma with MD diagnosis)

This is a 0/1 numeric variable. From the self-report on the respiratory history form, if the participant says that they ever had asthma and it was diagnosed by a doctor/HP, then:

If rsea32 is not missing:

If rsea32=0 or (rsea32=1 and rsea34=0) then asthma_ever_md=0.
Else if rsea32 =1 and rsea34=1 then asthma_ever_md=1.
Otherwise asthma_ever_md= missing.

Response format: 0 = No
1 = Yes

Source variable(s):

RSEA32. Ever had asthma

RSEA34. Asthma diagnosed by doctor/HP

6.3 Asthma_Curr (Currently have asthma)

This is a 0/1/9 numeric variable. From the self-report on the respiratory history form, if the participant says that they ever had asthma and if they still currently have asthma, then:

If RSEA32 is not missing and RSEA35 is not missing, then
If RSEA32 = 0 then ASTHMA_CURR= 0.
Else if RSEA32=1 and RSEA35=0 then ASTHMA_CURR= 0.
Else if RSEA32=1 and RSEA35=1 then ASTHMA_CURR= 1.
Else if RSEA32=1 and RSEA35 in (.,9) then ASTHMA_CURR= 9.

Response format: 0 = No
1 = Yes
9 = Don't Know

Source variable(s):

RSEA32. Ever had asthma

RSEA35. Still has asthma

6.4 Asthma_Curr_MD (Currently have asthma with ever MD diagnosis)

This is a 0/1 numeric variable. From the self-report on the respiratory history form, if the participant says that they ever had asthma that was diagnosed by a doctor/HP and if they still currently have asthma, then:

If Asthma_ever_md=0 and rsea35 not =1 or 9 then Asthma_Curr_MD=0

Else if Asthma_ever_md=0 and rsea35=1 then asthma_curr_md=0

Else if Asthma_ever_md=1 and rsea35=0 then Asthma_curr_md=0

Else if Asthma_Ever_MD is non-missing and rsea35=1 then Asthma_curr_md=1

Otherwise asthma_curr_md=.

Response format: 0 = No

1 = Yes

Source variable(s):

ASTHMA_EVER_MD. Ever had asthma with MD diagnosis

RSEA35. Still has asthma

6.5 Asthma_C4 (4-level grouped Asthma Status)

This is a 1/2/3/9 numeric variable. From the self-report on the respiratory history form, if the participant says that they ever had asthma or not and whether they still currently have asthma, then:

If ASTHMA_EVER is not missing then

If ASTHMA_EVER= 0 then ASTHMA_C4= 1.

Else if ASTHMA_EVER=1 and ASTHMA_CURR=0 then ASTHMA_C4= 2.

Else if ASTHMA_EVER =1 and ASTHMA_CURR =1 then ASTHMA_C4= 3.

Else if ASTHMA_EVER =1 and ASTHMA_CURR = missing or 9 then ASTHMA_C4= 9.

Response format: 1 = Never

2 = Former

3 = Current

9 = Don't Know

Source variable(s):

ASTHMA_EVER. Ever had asthma

ASTHMA_CURR. Still has asthma

6.6 Asthma_C4_MD (4-level grouped Asthma Status – by MD diagnosis)

This is a 1/2/3/9 numeric variable. From the self-report on the respiratory history form, if the participant says that they ever had asthma or not based on an MD diagnosis and whether they still currently have asthma, then:

If ASTHMA_EVER_MD is not missing then
If ASTHMA_EVER_MD= 0 then ASTHMA_C4_MD= 1.
Else if ASTHMA_EVER_MD =1 and ASTHMA_CURR_MD =0 then ASTHMA_C4_MD = 2.
Else if ASTHMA_EVER_MD =1 and ASTHMA_CURR_MD =1 then ASTHMA_C4_MD = 3.
Else if ASTHMA_EVER_MD =1 and ASTHMA_CURR_MD = missing or 9 then
ASTHMA_C4_MD = 9.

Response format: 1 = Never
2 = Former
3 = Current
9 = Don't Know

Source variable(s):

ASTHMA_EVER_MD. Ever had asthma based on MD diagnosis
ASTHMA_CURR_MD. Still has asthma based on ever MD diagnosis

6.7 COPD_Self (Self-report of ever had COPD/Emph)

From the self-report on the respiratory history form, if the participant says that they are told by a doctor that they have COPD or emphysema.

If RSEA44 is not missing, then COPD_Self=(RSEA44=1);

Source variable(s):

RSEA44. Ever had COPD or emphysema

6.8 COPD_EVER (Self-report of ever had COPD/Emph or CB)

From the self-report on the respiratory history form, if the participant says that they are told by a doctor that they have COPD or emphysema and/or chronic bronchitis.

If RSEA42 and RSEA44 is not missing, then.

If rsea42=0 and rsea44=0 then COPD_Ever=0
Else if rsea42=1 or rsea44=1 then COPD_EVER=1

Source variable(s):

RSEA42. Ever has chronic bronchitis
RSEA44. Ever had COPD or emphysema

6.9 MRC_CB (Chronic Productive Cough 3+ mths in 2+ yrs)

From the self-report on the respiratory history form, if the participant says that they had a cough on most days/nights of the week or when waking up in the morning, both for at least 3 months in a row, for at least the last 2 years, then MRC_CB is set to 1. If participant answered "no" to both having a cough on most days/nights of the week or when waking up in the morning, but answered that he/she had a cough for at least the last 2 years, then MRC_CB is still set to 1.

If rsea3 ≥ 2 years then rsea3_cat=1

Else if $0 \leq \text{rsea3} < 2$ years then $\text{rsea3_cat}=0$

If $\text{rsea1}=0$ and $\text{rsea2}=0$ and $\text{rsea3_cat} \neq 1$ then $\text{MRC_CB}=0$

Else if $\text{rsea3_cat}=0$ then $\text{MRC_CB}=0$

Else if $\text{rsea3_cat}=1$ then $\text{MRC_CB}=1$

5/2014 update: Count RSEA3=99 as missing

Source variable(s):

RSEA1. Ever had a cough on most days/nights of the week (at least 3 months in a row)

RSEA2. Ever had cough when waking up (at least 4 per week)

RSEA3. Years had cough (yes to rsea1 or rsea2 only)

6.10 Wheeze_SB (Wheezing ever with shortness of breath)

From the self-report on the respiratory history form, if the participant says that they had wheezing or whistling in chest and if they ever had shortness of breath from wheezing or whistling in chest, then:

If $\text{rsea7}=0$ then $\text{Wheeze_SB}=0$

Else if $\text{rsea9}=0$ then $\text{Wheeze_SB}=0$

Else if $\text{rsea9}=1$ then $\text{Wheeze_SB}=1$

Source variable(s):

RSEA7. Ever had wheezing or whistling in chest

RSEA9. Ever had shortness of breath from wheezing or whistling in chest

6.11 Wheeze_Last12 (Wheezing in last 12 months)

From the self-report on the respiratory history form, if the participant says that they had wheezing or whistling in chest and if they had it in the last 12 months, then:

If $\text{rsea7}=0$ then $\text{Wheeze_Last12}=0$

Else if $\text{rsea13}=0$ then $\text{Wheeze_Last12}=0$

Else if $\text{rsea13}=1$ then $\text{Wheeze_Last12}=1$

Source variable(s):

RSEA7. Ever had wheezing or whistling in chest

RSEA13. Ever had wheezing or whistling in chest in the last 12 months

6.12 Phlegm (Phlegm day/night 3 months in a row)

From the self-report on the respiratory history form, if the participant says that they had phlegm in chest most days or nights and if they had it for at least 3 months in a row, then:

If $\text{rsea4}=0$ then $\text{Phlegm}=0$

Else if $\text{resa4}=1$ then $\text{Phlegm}=1$

Else $\text{Phlegm}=\text{missing}$

Source variable(s):

RSEA4. Phlegm from chest on most days/nights of the week during at least three months in a row

6.13 Pneumonia (Pneumonia - by MD diagnosis)

From the self-report on the respiratory history form, if the participant says that a doctor has ever told them that they had pneumonia or bronchopneumonia, then:

If rsea40=0 then Pneumonia=0
Else if resa40=1 then Pneumonia=1
Else Pneumonia=missing

Source variable(s):

RSEA40. Ever have pneumonia or bronchopneumonia based on MD diagnosis

6.14 Valid_spirometry (Valid spirometry studies, pre-BD)

This 0/1 variable is an overall data quality indicator derived by the pulmonary reading center on the adequacy study for use. The value is 0 for not usable vs. 1 for usable for analysis.

Response format: 0 = No acceptable curves for either FEV₁ or FVC quality attribute
1 = At least one acceptable curve for both FEV₁ or FVC quality attributes

Source variable(s):

PRBA3. FVC quality attribute (pre-BD)
PRBA4. FEV₁ quality attribute (pre-BD)

6.15 Valid_spirometry_postBD (Valid spirometry studies, post-BD)

This variable is an overall data quality indicator derived by the pulmonary reading center on the adequacy study for use. The value is 0 for not usable vs. 1 for usable for analysis.

Response format: 0 = No acceptable curves for either FEV₁ or FVC quality attribute
1 = At least one acceptable curve for both FEV₁ or FVC quality attributes

Source variable(s):

POBA3. FVC quality attribute (post-BD)
POBA4. FEV₁ quality attribute (post-BD)

6.16 Airflow_Obstruction (Obstructive Lung Disease)

This is a 0/1 variable based on the empirical evidence from the spirometry that the ratio of FEV₁ to FVC is below the level defined by the ATS as indicative of restricted airflow.

If PRBA25 is not missing (and not equal to 0) and PRBA29 is not missing then
AIRFLOW_OBSTRUCTION= 1 if
(PRBA29 / PRBA25) < 0.70 or (PRBA29 / PRBA25) *100 < prba46.

Else AIRFLOW_OBSTRUCTION= 0.

5/2014 update: set to missing for invalid spirometries

Source variable(s):

PRBA25. FVC (mL)

PRBA29. FEV₁ (mL)

PRBA46. Lower Limit of Normal FEV1/FVC%

6.17 Airflow_Limitation (Airflow Limitation (pre FEV1/FVC<70 only))

This is a 0/1 variable based on the empirical evidence from the spirometry that the ratio of FEV₁ to FVC is below the level defined by the ATS as indicative of restricted airflow. Defined only for valid spirometries.

If PRBA25 is not missing and PRBA29 is not missing then
AIRFLOW_LIMITATION= 1 if $(PRBA29 / PRBA25) < 0.70$.
Else AIRFLOW_LIMITATION= 0.

Source variable(s):

PRBA25. FVC (mL)

PRBA29. FEV₁ (mL)

Valid_spirometry. Valid spirometry studies, pre-BD

6.18 COPD_by_BD (COPD by bronchodilator test)

This indicator of COPD is based on the post-bronchodilator measurement of the FEV₁ to FVC ratio. The airflow remains restricted after the administration of albuterol.

If POBA25 is not missing (or zero) and POBA29 is not missing then COPD_BY_BD = $(POBA29 / POBA25) < 0.70$;

Response format: 0 = No
1 = Yes

Source variable(s):

POBA25. FVC (mL)

POBA29. FEV₁ (mL)

6.19 COPD_by_BD_AllPre (COPD by BD test (for all valid pre_BD))

This is a 0/1 variable based on the post-bronchodilator measurement of the FEV₁ to FVC ratio. The airflow remains restricted after the administration of albuterol. This variable is also defined for all participants who did not conduct a post-spirometry exam. Defined only for valid spirometries

If valid_spirometry=0 then do the following:
COPD_BY_BD_ALLPRE=.
Else if valid_spirometry=1 & POSTBD_EXPECTED=0 then set
COPD_BY_BD_ALLPRE=0

Else if valid_spirometry_postbd2=2 then set COPD_BY_BD_ALLPRE=
Otherwise set COPD_BY_BD_ALLPRE=COPD_BY_BD

5/2014 update: set to missing for invalid spirometries

Source variable(s):

COPD_by_BD. COPD by bronchodilator test

VALID_SPIROMETRY. Valid spirometry studies (pre BD).

Valid_spirometry_postbd2. Valid spirometry (A-D grading) adjusted for unexpected BDs.

Postbd_expected. Post-BD expected among valid spirometry.

6.20 BD_response_abs (Bronchodilator Responsive by 15% change)

This is a 0/1 variable based on the post-bronchodilator measurement of the FEV₁. The airflow remains restricted after the administration of albuterol the person did not respond by greater than a 15% improvement over their first measurement.

If PRBA29 is not missing and POBA29 is not missing and PRBA29 > 0 then

If (((poba29-prba29)/prba29)>0.15) then BD_response_abs= 1.

Otherwise if $0 \leq ((poba29-prba29)/prba29) \leq 0.15$ then BD_Response_abs=0.

Response format: 0 = less than or equal to 15% improvement

1 = greater than 15% improvement

Source variable(s):

PRBA29. FEV₁ (mL) pre-BD measurement

POBA29. FEV₁ (mL)

6.21 FEV1_FVC_Ratio (FEV₁ to FVC %)

This variable is the ratio of forced expiratory volume in one minute (FEV₁) to forced vital capacity (FVC) expressed as a percentage.

Only defined for VALID_SPIROMETRY=1, else missing

If PRBA25 is not missing (and not equal to 0) and PRBA29 is not missing then

$$\text{FEV1_FVC_Ratio} = (\text{PRBA29} / \text{PRBA25}) * 100$$

5/2014 update: set to missing for invalid spirometries

Source variable(s):

VALID_SPIROMETRY. Valid spirometry studies (pre BD).

VALID_SPIROMETRY_POSTBD (Valid spirometry studies, post-BD)

PRBA25. FVC (mL)

PRBA29. FEV₁ (mL)

6.22 POSTBD_PREBD_DIFF (Post(FEV1%pred)-Pre(FEV1%pred))

This variable is based on the ratio of the measured FEV₁ difference from post-bronchodilator and pre-bronchodilator exams over the Predicted FEV₁ calculated pre-

bronch (both values must be non-missing, and the post-bronchodilator spirometry examinations must be valid), as follows:

Only defined if valid_spirometry=1 and valid_spirometry_postbd=1, else missing
PostBD_PreBD_Diff = ((POBA29-PRBA29)/PRBA37)*100.

5/2014 update: set to missing for invalid spirometries

Source variable(s):

VALID_SPIROMETRY. Valid spirometry studies (pre BD).

PRBA29. FEV₁ (mL) pre-BD measurement

PRBA37. Predicted FEV₁ (mL) (recalculated)

POBA29. FEV₁ (mL) post-BD measurement

6.23 POSTBD_PREBD_DIFF_GT12 (Relative Bronchodilator Responsive by 12% change)

This is a 0/1 variable based on the ratio of the measured FEV₁ difference from post-bronchodilator and pre-bronchodilator exams over the Predicted FEV₁ calculated pre-bronch (both values must be non-missing, and the post-bronchodilator spirometry examination must be valid) as follows:

if PostBD_PreBD_Diff \geq 12 then PostBD_PreBD_Diff_GT12_2 = 1.

else if $0 \leq$ PostBD_PreBD_Diff < 12 then PostBD_PreBD_Diff_GT12_2 = 0.

Source variable(s):

POSTBD_PREBD_DIFF. Post(FEV1%pred)-Pre(FEV1%pred)

FEV1_FVC_Ratio. FEV₁ to FVC %

PRBA46. Lower Limit of Normal FEV1/FVC% (recalculated)

Valid_spirometry. Valid spirometry studies, pre-BD

6.24 POSTBD_PREBD_DIFF_GT15 (Relative Bronchodilator Responsive by 15% change)

This is a 0/1 variable based on the ratio of the measured FEV₁ difference from post-bronchodilator and pre-bronchodilator exams over the Predicted FEV₁ calculated pre-bronch (both values must be non-missing, and the post-bronchodilator spirometry examination must be valid) as follows:

if PostBD_PreBD_Diff \geq 15 then PostBD_PreBD_Diff_GT12_2 = 1.

else if $0 \leq$ PostBD_PreBD_Diff < 15 then PostBD_PreBD_Diff_GT12_2 = 0.

Source variable(s):

POSTBD_PREBD_DIFF. Post(FEV1%pred)-Pre(FEV1%pred)

FEV1_FVC_Ratio. FEV₁ to FVC %

PRBA46. Lower Limit of Normal FEV1/FVC% (recalculated)

Valid_spirometry. Valid spirometry studies, pre-BD

6.26 POSTBD_PREBD_DIFF_GT12_2 (Relative Bronchodilator Responsive by 12% change (among pre FEV1/FVC<70 only))

This is a 0/1 variable based on the ratio of the measured FEV₁ difference from post-bronchodilator and pre-bronchodilator exams over the Predicted FEV₁ calculated pre-bronch (both values must be non-missing, and the post-bronchodilator spirometry examination must be valid) **among participants with pre FEV1/FVC<70 only (ignores LLN comparison)**, as follows:

if PostBD_PreBD_Diff \geq 12 then PostBD_PreBD_Diff_GT12_2 = 1.
else if $0 \leq$ PostBD_PreBD_Diff < 12 then PostBD_PreBD_Diff_GT12_2 = 0.

Source variable(s):

POSTBD_PREBD_DIFF. Post(FEV1%pred)-Pre(FEV1%pred)
FEV1_FVC_Ratio. FEV₁ to FVC %
PRBA46. Lower Limit of Normal FEV1/FVC% (recalculated)
Valid_spirometry. Valid spirometry studies, pre-BD

6.27 POSTBD_EXPECTED (Post-BD Expected (among valid spirometry))

This variable is a flag for whether a post bronchodilator was expected, based on the pre-measurement. A post measurement was expected if the FEV₁ to FVC ratio was less than 70% or less than the lower limit of normal for FEV₁/FVC%. Note the flag is missing for participants who did not have a valid pre measurement.

if VALID_SPIROMETRY=1 then do the following:

If $0 \leq$ (FEV1_FVC_RATIO/100) < 0.70 or $0 \leq$ FEV1_FVC_RATIO < PRBA46 then set POSTBD_EXPECTED=1

Otherwise, set POSTBD_EXPECTED=0.

Response format: 0=Not Expected
1= Expected

Source variable(s):

VALID_SPIROMETRY. Valid Spirometry Studies (Pre)
FEV1_FVC_RATIO. FEV₁ to FVC Ratio (%)
PRBA46. Lower Limit of Normal FEV1/FVC% (recalculated)

6.28 VALID_SPIROMETRY_POSTBD2 (Valid post-spirometry (A-D grading) adjusted for unexpected BDs)

This variable is a flag for whether a post bronchodilator was valid (grade A-D). Defined only for participants where a post measurement was expected.

if POSTBD_EXPECTED=1 then set VALID_SPIROMETRY_POSTBD2 =
VALID_SPIROMETRY_POSTBD

Response format: 0=Invalid
1=Valid

Source variable(s):

POSTBD_EXPECTED. Post-BD Expected (among valid spirometry)
VALID_SPIROMETRY_POSTBD. Valid Spirometry Studies (Post BD)

7. WELL-BEING/QUALITY OF LIFE

7.1 CESD10 (10-item CES-D summary score)

This is a numeric variable with values ranging from 0 to 30.

Missing if more than 2 of the 10 items (WBEA1-WBEA10) are missing

Sum of (WBEA1-WBEA10), where positively worded items (WBEA5, WBEA8) are reverse coded. For participants with 1 or 2 missing items, CESD10 is equal to the weighted average of the nonmissing items*10

5/2014 update: for participants with 1 or 2 missing items, the weighted average of the nonmissing items*10 is used instead of the sum of the nonmissing items

Response format for WBEA1-WBEA10:

- 0 = rarely or none of the time
- 1 = some or a little of the time
- 2 = occasionally or a moderate amount of time
- 3 = all of the time

Source variable(s):

WBEA1. I was bothered by things that usually don't bother me

WBEA2. I had trouble keeping my mind on what I was doing

WBEA3. I felt depressed

WBEA4. I felt that everything I did was an effort

WBEA5. I felt hopeful about the future

WBEA6. I felt fearful

WBEA7. My sleep was restless

WBEA8. I was happy

WBEA9. I felt lonely

WBEA10. I could not "get going"

7.2 STAI10 (10-item State-Trait Anxiety Inventory summary score)

This is a numeric variable with values ranging from 0 to 40.

Missing if more than 2 of the 10 items (WBEA11-WBEA20) are missing

Sum of (WBEA11-WBEA20), where positively worded items (WBEA12, WBEA17, WBEA19) are reverse coded, and 1 is added to all responses (converting from 0-3 to 1-4 scale to match the standardized instrument scoring). For participants with 1 or 2 missing items, STAI10 is equal to the weighted average of the nonmissing items*10

5/2014 update: for participants with 1 or 2 missing items, the weighted average of the nonmissing items*10 is used instead of the sum of the nonmissing items

Response format for WBEA11-WBEA20:

0 = almost never
1 = sometimes
2 = often
3 = always

Source variable(s):

WBEA11. I feel nervous and restless.

WBEA12. I feel satisfied with myself.

WBEA13. I wish I could be as happy as others seem to be.

WBEA14. I feel like a failure.

WBEA15. I worry too much over something that really doesn't matter.

WBEA16. I lack self-confidence.

WBEA17. I feel secure.

WBEA18. I feel inadequate.

WBEA19. I am a steady person.

WBEA20. I get in a state of tension or turmoil as I think over my recent concerns and interests.

8. SF-12v2 HEALTH SURVEY

8.1 Agg_ment (Aggregate Mental Health Scale)

This is a numeric variable. The factor scores and general population norms developed by John Ware for scoring the SF-12v2© (see pp. 45-51 of “How to Score Version 2 of the SF-12 Health Survey, QualityMetrics, 2002) are used to compute this summary measure for the mental health domain of quality of life. This sub-scale score is a norm-based transformation of standardized Z scores of the constituent items for the established domains which is scaled to a mean of 50 and standard deviation of 10.

8.2 Agg_phys (Aggregate Physical Health Scale)

This is a numeric variable. The factor scores and general population norms developed by John Ware for scoring the SF-12v2© (see pp. 45-51 of “How to Score Version 2 of the SF-12 Health Survey, QualityMetrics, 2002) are used to compute this summary measure for the physical health domain of quality of life. This sub-scale score is a norm-based on transformation of a standardized Z scores of the constituent items which is scaled to a mean of 50 and standard deviation of 10.

9. BLOOD PRESSURE MEASURES

9.1 Left_ABI, Right_ABI, and ABI_Overall (Ankle Brachial Index –occ. failure included)

This is a numeric variable. Ankle Brachial Index (ABI) is the average ankle to arm systolic pressures as measured from the left and right sides. The higher of two ankle measures on the same side is divided by the higher of the two systolic arm pressures to derive the right and left side ABI values. The overall ABI is the mean of those two measurements.

Occlusion failures: missing observation values were set to 300 where occlusion failures occurred as follows:

If ABPA1B was missing and ABPA3A2=1 then ABPA1B was set to 300

If ABPA1C was missing and ABPA3A3=1 then ABPA1C was set to 300

If ABPA1D was missing and ABPA3A4=1 then ABPA1D was set to 300

If ABPA1E was missing and ABPA3A5=1 then ABPA1E was set to 300

There are were 2 cases where ABI variables were set to missing:

1. If both ABPA1A and ABPA1F were missing.

2. All three of the variables on the same side of the body are =0 as follows:

ABPA1A, ABPA1B, and ABPA1C

ABPA1D, ABPA1E, and ABPA1F

Left side Ankle Brachial Index:

$LEFT_ABI = \text{maximum of (abpa1d, abpa1e)} / \text{maximum of (abpa1a, abpa1f)}$

Right side Ankle Brachial Index:

$RIGHT_ABI = \text{maximum of (abpa1b, abpa1c)} / \text{maximum of (abpa1a, abpa1f)}$

Overall Ankle Brachial Index:

$ABI_OVERALL = \text{mean of (left_ABI, right_ABI)}$

5/2014 Update: ABI definitions updated to include occlusion failures

Source variable(s):

ABPA1A. Right brachial

ABPA1B. Right dorsalis pedis

ABPA1C. Right posterior tibial

ABPA1D. Left posterior tibial

ABPA1E. Left dorsalis pedis

ABPA1F. Left brachial

ABPA3A2. Occlusion failure of R. dorsalis pedis

ABPA3A3. Occlusion failure of R. posterior tibial

ABPA3A4. Occlusion failure of L. posterior tibial

ABPA3A5. Occlusion failure of L. dorsalis pedis

9.2 LEFT_ABIGRP_C4, RIGHT_ABIGRP_C4 (4-level grouped Ankle Brachial Index (occ. failure incl.))

This variable groups continuous left and right ABI into 4 categories:

When $0 \leq \text{left_ABI} \leq 0.90$ then set $\text{left_ABIGRP_C4} = 1$
When $0.90 < \text{left_ABI} < 1.0$ then set $\text{left_ABIGRP_C4} = 2$
When $1.0 \leq \text{left_ABI} < 1.4$ then set $\text{left_ABIGRP_C4} = 3$
Otherwise when $\text{left_ABI} \geq 1.40$ then set $\text{left_ABIGRP_C4} = 4$

RIGHT_ABIGRP_C4 is defined similarly.

Response format: 1=Low (PAD)
2=Borderline
3=Normal
4=High

Source variable(s):

LEFT_ABI. Right Ankle Brachial Index (occ. failure incl.)
RIGHT_ABI. Left Ankle Brachial Index (occ. failure incl.)

9.3 ABIGRP_C4 (4-level grouped Ankle Brachial Index (occ. failure incl.))

This variable is based on the calculated ankle brachial index (mean of left/right) as follows:

if $0 \leq \text{ABI_OVERALL} \leq 0.90$ then $\text{ABIGRP_C4} = 1$;
else if $0.90 < \text{ABI_OVERALL} < 1.0$ then $\text{ABIGRP_C4} = 2$;
else if $1.0 \leq \text{ABI_OVERALL} < 1.4$ then $\text{ABIGRP_C4} = 3$;
else if $\text{ABI_OVERALL} \geq 1.40$ then $\text{ABIGRP_C4} = 4$;

Response format: 1=Low (PAD)
2=Borderline
3=Normal
4=High

Source variable(s):

ABI_OVERALL. Overall Ankle Brachial Index (occ. failure incl.)

9.4 Left_ABI2, Right_ABI2, and ABI2_Overall (Ankle Brachial Index - version 2)

This is a numeric variable based on calculation methods from 2010 (AIE 2010; 171:368–376). Ankle Brachial Index (ABI) is the average ankle to arm systolic pressures as measured from the left and right sides. The higher of two ankle measures on the same side is divided by either the average of the two systolic arm pressures (if they differ by less than 15 mmHg) or the higher of the two systolic arm pressures (if they differ by more than 15 mmHg) to derive the right and left side ABI values. The overall ABI is the minimum of those two measurements.

If absolute value of (abpa1a - abpa1f) > 15 then:

Left side Ankle Brachial Index:

LEFT_ABI_v2=maximum of (abpa1d,abpa1e) / maximum of (abpa1a,abpa1f)

Right side Ankle Brachial Index:

RIGHT_ABI_v2=max(abpa1b,abpa1c)/maximum of (abpa1a,abpa1f)

If absolute value of (abpa1a - abpa1f) ≤ 15 then:

Left side Ankle Brachial Index:

LEFT_ABI_v2=maximum of (abpa1d,abpa1e) / mean of (abpa1a,abpa1f)

Right side Ankle Brachial Index:

RIGHT_ABI_v2=maximum of (abpa1b,abpa1c)/mean of (abpa1a,abpa1f)

Overall Ankle Brachial Index:

ABI_OVERALL_v2=min of (left_ABI_v2, right_ABI_v2)

Source variable(s):

ABPA1A. Right brachial

ABPA1B. Right dorsalis pedis

ABPA1C. Right posterior tibial

ABPA1D. Left posterior tibial

ABPA1E. Left dorsalis pedis

ABPA1F. Left brachial

9.5 ABIGRP2_C4 (4-level grouped Ankle Brachial Index - version 2)

This variable is based on the calculated ankle brachial index (min of left/right) as follows:

if $0 \leq \text{ABI_OVERALL_v2} \leq 0.90$ then ABIGRP2_C4=1;
else if $0.90 < \text{ABI_OVERALL_v2} < 1.0$ then ABIGRP2_C4=2;
else if $1.0 \leq \text{ABI_OVERALL_v2} < 1.4$ then ABIGRP2_C4=3;
else if $\text{ABI_OVERALL_v2} \geq 1.40$ then ABIGRP2_C4=4;

Response format: 1=Low (PAD)

2=Borderline

3=Normal

4=High

Source variable(s):

ABI_OVERALL_v2. Ankle Brachial Index – v2.

10. LABORATORY MEASURES

10.1 CKD (Chronic Kidney Disease using eGFR only - NIDDK)

This variable is estimated based on the estimated GFR value and is defined using NIDDK guidelines, as follows:

if $GFR \geq 90$ then $CKD=1$ ('Normal')
else if $60 \leq GFR < 90$ then $CKD=2$ ('Mild')
else if $30 \leq GFR < 60$ then $CKD=3$ ('Moderate')
else if $15 \leq GFR < 30$ then $CKD=4$ ('Severe')
else if $.Z < GFR < 15$ then $CKD=5$ ('End-Stage')

Response format: 1=Normal
2=Mild
3=Moderate
4=Severe
5=End-Stage

Source variable(s):

GFR. Glomerular Filtration Rate (mL/min/1.73m²), estimated using the Modification of Diet in Renal Disease (MDRD) Study equation¹, which can be found at <https://www.niddk.nih.gov/health-information/health-communication-programs/nkdep/lab-evaluation/gfr-calculators/adults-conventional-unit/Pages/adults-conventional-unit.aspx#2>

Specifically, the MDRD study equation is:

$GFR \text{ (mL/min/1.73 m}^2\text{)} = 175 \times (S_{cr})^{-1.154} \times (\text{Age})^{-0.203} \times (0.742 \text{ if female}) \times (1.212 \text{ if African American})$
(conventional units)

where S_{cr} is serum creatinine (LABA76)

¹Levey AS, Coresh J, Greene T, Stevens LA, Zhang YL, Hendriksen S, Kusek JW, Van Lente F, Chronic Kidney Disease Epidemiology Collaboration. Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate. *Ann Intern Med.* 2006;145:247–254.

10.2 Insulin_Fast (Insulin, fasting (calibrated, converted to mU/L))

This variable combines fasting glucose results from the central lab that was processed prior to July 5, 2010 (using units of mU/L) with results processed on or after July 5, 2010 (using units of pmol/L) and the overall units are converted to mU/L (1 mU/L = 6 pmol/L).

In addition, values collected prior to October 29, 2009 have been calibrated using the following regression equation: $y = 1.00494x - 1.4504$

Where y = adjusted insulin value, x = original insulin value using Merdcodia assay.

By doing so, there is no overlap in the adjustment for a change in assay from Merdcodia to Roche Elecsys analysers which occurred on October 29, 2009. Update 12/2013: changed missing to .L for .L values of LABA73

Source variable(s):

LABA73. Fasting Insulin (mU/L)

LABA96. Fasting Insulin (pmol/L)

10.3 Insulin_OGTT (Insulin, post OGTT (calibrated, converted to mU/L))

This variable combines post OGTT glucose results from the central lab that were processed prior to July 5, 2010 (using units of mU/L) with results processed on or after July 5, 2010 (using units of pmol/L) and the overall units are converted to mU/L (1 mU/L = 6 pmol/L).

In addition, values collected prior to October 29, 2009 have been calibrated using the following regression equation: $y = 1.00494x - 1.4504$

Where y = adjusted insulin value, x = original insulin value using Merdcodia assay.

By doing so, there is no overlap in the adjustment for a change in assay from Merdcodia to Roche Elecsys analysers which occurred on October 29, 2009.

5/2014 update: corrected to account for .N values of insulin

Source variable(s):

LABA90. Post OGTT Insulin (mU/L)

LABA97. Post OGTT Insulin (pmol/L)

10.4 Insulin_Date (Date of Insulin Measures - All)

This variable combines date variables for insulin testing from the central lab that were processed prior to July 5, 2010 with those processed on or after July 5, 2010.

Source variable(s):

LABA73A. Date of Insulin measures (before July 5, 2010)

LABA96A. Date of Insulin measures (on or after July 5, 2010)

10.5 Dyslipidemia (High LDL, low HDL, or high triglycerides - NIH)

This is a 0/1 variable that checks for HDL, LDL, and triglyceride values to determine presence/absence of dyslipidemia as follows:

If $laba69 \leq .Z$ & $laba68 \leq .Z$ & $laba67 \leq .Z$ then DYSLIPIDEMIA= missing.

Else if $laba69 \geq 160$ OR $0 \leq laba68 < 40$ OR $laba67 \geq 200$ then DYSLIPIDEMIA=1

Else DYSLIPIDEMIA=0.

Response format: 0= Not dyslipidemic (No)

1= Dyslipidemic (Yes)

Source variable(s):

LABA67. Triglycerides (mg/dL)

LABA68. HDL - cholesterol (mg/dL)

LABA69. LDL - Cholesterol (mg/dL)

10.6 Dyslipidemia_C3 (3-Category Dyslipidemia)

This is a numeric variable that describes the status of dyslipidemia based on NIH definition and uses of self-reported or scanned medication.

```
if dyslipidemia<=.Z & med_lld^=1 & muea33e^=1 then DYSLIPIDEMIA_C3=.;  
else if dyslipidemia=0 & med_lld^=1 & muea33e^=1 then DYSLIPIDEMIA_C3=1;  
else if dyslipidemia=1 & med_lld^=1 & muea33e^=1 then DYSLIPIDEMIA_C3=2;  
else if dyslipidemia=1 | med_lld=1 | muea33e=1 then DYSLIPIDEMIA_C3=3;
```

Response format: 1 = Non-dyslipidemia
2 = Dyslipidemia w/o Treatment
3 = Dyslipidemia with Treatment

Source variable(s):

Dyslipidemia. High LDL, low HDL, or high triglycerides - NIH
MED_LLD. Scanned Lipid lowering drugs/Antihyperlipidemics (LLD)
MUEA33e. Self-reported Antihyperlipidemics medicación used.

10.7 Cardiac_Risk_Ratio (Cardiac Risk Ratio (TC/HDL) - AHA)

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

If laba66 and laba68 are both non-missing, then:
CARDIAC_RISK_RATIO=laba66/laba68.

Source variable(s):

LABA66. Total Cholesterol (mg/dL)
LABA68. HDL - cholesterol (mg/dL)

10.8 Dyslipidemia_TCHDL (Dyslipidemia based on Cardiac Risk Ratio > 5)

This is a 0/1 variable that checks for Dyslipidemia based on Cardiac Risk Ratio (TC/HDL) > 5 as follows:

```
if cardiac_risk_ratio<=.Z then DYSLIPIDEMIA_TCHDL=.;  
else if cardiac_risk_ratio>5 then DYSLIPIDEMIA_TCHDL=1;  
else DYSLIPIDEMIA_TCHDL=0;
```

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

Source variable(s):

Cardiac_Risk_Ratio. Cardiac_Risk_Ratio (TC/HDL) AHA

10.9 DYS_TCHDL_MED (Dyslipidemia based on Cardiac Risk Ratio > 5 or Lipid Lower Drug scanned at baseline)

This is a 0/1 variable that checks for Dyslipidemia defined by Cardiac Risk Ratio (TC/HDL) > 5 or Lipid Lower Drug scanned at baseline as follows:

if dyslipidemia_tchdl<=.Z & med_llid<=.Z then DYS_TCHDL_MED=.;
else if dyslipidemia_tchdl=1 | med_llid=1 then DYS_TCHDL_MED=1;
else DYS_TCHDL_MED=0;

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

Source variable(s):

Cardiac_Risk_Ratio. Cardiac_Risk_Ratio (TC/HDL) AHA
MED_LLD. Lipid lowering drugs/Antihyperlipidemics (LLD)

10.10 High_total_Chol (Hypercholesterolemia (from LABA66, LABA68, LABA69, MED_LLD))

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

If missing(laba66) & missing(laba68) & missing(laba69) and missing(med_llid) then
High_total_Chol=.;
Else if (laba66>=240) OR (laba69>= 160) OR (0<laba68<40) OR (MED_LLD=1)
then High_total_Chol =1;
Else High_total_Chol =0;

5/2014 update: updates due to updated medication variables

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

Source variable(s):

LABA66. Total Cholesterol (mg/dL)
LABA68. HDL - cholesterol (mg/dL)
LABA69. LDL - Cholesterol (mg/dL)
MED_LLD. Lipid lowering drugs/Antihyperlipidemics (LLD)

10.11 TOTAL_CHOL_C3 (3-Category Total Cholesterol)

This is a numeric variable that describes the stages of total cholesterol based on lab measurement as follow.

if .Z<laba66<200 then TOTAL_CHOL_C3=1;
else if 200<=laba66<=239 then TOTAL_CHOL_C3=2;
else if laba66>=240 then TOTAL_CHOL_C3=3;
else TOTAL_CHOL_C3=.;

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

Source variable(s):

LABA66. Total Cholesterol (mg/dL)

10.12 Fasttime (Fasting time in hours)

Elapsed time between the time the participant last consumed anything and the blood draw. This is a numeric variable scaled in hours.

Source variable(s):

BIOA7. Day last meal = (1= today, 0hr. adjustment; 2=yesterday, 12hr. adjustment; 3=day before yesterday, 24hr. adjustment.)

BIOA8. time of last meal

BIOA9. date of blood draw

BIOA10. time of blood draw

10.13 IFG (Impaired Fasting Glucose - AHA)

This is a 0/1 variable that checks for Impaired Fasting Glucose by AHA Guideline as follows:

if $100 \leq \text{LABA70} \leq 125$ & (MED_ANTIDIAB NE 1) then IFG=1;
else if $0 < \text{LABA70} < 100$ | $\text{LABA70} > 125$ then IFG=0;

Response format: 0= Not Impaired Fasting Glucose (No)
1= Impaired Fasting Glucose (Yes)

Source variable(s):

LABA70. Fasting glucose (mg/dL)

MED_ANTIDIAB. Scanned antidiabetics medication

10.14 IGT (Impaired Glucose Tolerance - AHA)

This is a 0/1 variable that checks for Impaired Glucose Tolerance by AHA Guideline as follows:

if $0 < \text{LABA70} < 126$ & $140 \leq \text{laba71} < 200$ then IGT=1;
else if $\text{LABA70} > .Z$ & $\text{LABA71} > .Z$ then IGT=0;

Response format: 0= Not Impaired Glucose Tolerance (No)
1= Impaired Glucose Tolerance (Yes)

Source variable(s):

LABA70. Fasting glucose (mg/dL)

LABA71. Glucose, post OGGT (mg/dL)

10.15 GFR (Glomerular Filtration Rate, estimated)

This variable is estimated based on standard reference equations for Hispanics provided by the central clinical laboratory. The GFR value has been adjusted for age and gender.

10.16 GFRscys (eGFR based on serum cystatin C w/o demographics)

Estimated Glomerular filtration rate, based on serum cystatin C and no demographic factors.

Reference:

Estimating GFR using serum cystatin C alone and in combination with serum creatinine: a pooled analysis of 3,418 individuals with CKD. Am J Kidney Dis. 2008 Mar ;51(3):395-406

$$\text{GFRscys} = 76.7 * \text{LABA101}^{-1.19}$$

Source variable(s):

LABA101. Cystatin C (mg/L)

10.17 GFRscys_scr (eGFR based on serum cystatin C, serum creatinine, gender, age and race)

Estimated Glomerular filtration rate, based on serum cystatin C, serum creatinine, and demographics gender, age and race (black vs. nonblack).

Reference:

Estimating glomerular filtration rate from serum creatinine and cystatin C. N Engl J Med. 2012 Jul 5 ;367(1):20-9

If (GENDER=F and RACE=4) then $135 * \min(\text{LABA76}/0.7, 1)^{-0.248} * \max(\text{LABA76}/0.7, 1)^{-0.601} * \min(\text{LABA101}/0.8, 1)^{-0.375} * \max(\text{LABA101}/0.8, 1)^{-0.711} * 0.995^{\text{AGE}} * 0.969 * 1.08$

If (GENDER=F and RACE≠4) then $135 * \min(\text{LABA76}/0.7, 1)^{-0.248} * \max(\text{LABA76}/0.7, 1)^{-0.601} * \min(\text{LABA101}/0.8, 1)^{-0.375} * \max(\text{LABA101}/0.8, 1)^{-0.711} * 0.995^{\text{AGE}} * 0.969$

If (GENDER=M and RACE=4) then $135 * \min(\text{LABA76}/0.9, 1)^{-0.207} * \max(\text{LABA76}/0.9, 1)^{-0.601} * \min(\text{LABA101}/0.8, 1)^{-0.375} * \max(\text{LABA101}/0.8, 1)^{-0.711} * 0.995^{\text{AGE}} * 1.08$

If (GENDER=M and RACE≠4) then $135 * \min(\text{LABA76}/0.9, 1)^{-0.207} * \max(\text{LABA76}/0.9, 1)^{-0.601} * \min(\text{LABA101}/0.8, 1)^{-0.375} * \max(\text{LABA101}/0.8, 1)^{-0.711} * 0.995^{\text{AGE}}$

Source variable(s):

GENDER. Gender

RACE. Race Group (self-report)

AGE. Age (in years at the time of participant's clinic visit)

LABA76. Creatinine (mg/dL)

LABA101. Cystatin C (mg/L)

10.18 GFRscr (eGFR based on serum creatinine, gender, age, and race)

Estimated Glomerular filtration rate, based on serum creatinine, and demographics gender, age and race (black vs. nonblack).

Reference:

Estimating glomerular filtration rate from serum creatinine and cystatin C. N Engl J Med. 2012 Jul 5 ;367(1):20-9

If (GENDER=F and RACE=4) then $141 * \min(\text{LABA76}/0.7, 1)^{-0.329} * \max(\text{LABA76}/0.7, 1)^{-1.209} * 0.993^{\text{AGE}} * 1.018 * 1.159$

If (GENDER=F and RACE≠4) then $141 * \min(\text{LABA76}/0.7, 1)^{-0.329} * \max(\text{LABA76}/0.7, 1)^{-1.209} * 0.993^{\text{AGE}} * 1.018$

If (GENDER=M and RACE=4) then $141 * \min(\text{LABA76}/0.9, 1)^{-0.411} * \max(\text{LABA76}/0.9, 1)^{-1.209} * 0.993^{\text{AGE}} * 1.159$

If (GENDER=M and RACE≠4) then $141 * \min(\text{LABA76}/0.9, 1)^{-0.411} * \max(\text{LABA76}/0.9, 1)^{-1.209} * 0.993^{\text{AGE}}$

Source variable(s):

GENDER. Gender

RACE. Race Group (self-report)

AGE. Age (in years at the time of participant's clinic visit)

LABA76. Creatinine (mg/dL)

11. GLOBAL PHYSICAL ACTIVITY QUESTIONNAIRE (GPAQ)

The GPAQ was developed by the World Health Organization (WHO). It collects information on physical activity participation in three settings (or domains: work, travel, leisure) and sedentary behavior. Scoring information, available at www.who.int/chp/steps, was used to define the variables below.

A MET (Metabolic Equivalent) is the ratio of the metabolic rate for a given activity relative to resting metabolic rate. One MET is defined as 1 kcal/kg/hour and is equivalent to the energy cost of sitting quietly. A MET is also defined as oxygen uptake in ml/kg/min with one MET equal to the oxygen cost of sitting quietly, around 3.5 ml/kg/min.

For calculations described below, MET values were assigned value corresponding to an average MET for typical types and intensities of activity.

Domain	METS value
Work	<ul style="list-style-type: none">Moderate MET value = 4.0Vigorous MET value = 8.0
Transport	<ul style="list-style-type: none">Cycling and walking MET value = 4.0
Recreation	<ul style="list-style-type: none">Moderate MET value = 4.0Vigorous MET value = 8.0

Analyses information can be found in the following WHO documentation:
http://www.who.int/chp/steps/resources/GPAQ_Analysis_Guide.pdf

11.1 GPAQ_rec_mod (Recreational physical (moderate) activity (min/day))

This is a numeric variable that calculates the average amount of time spent per day doing a moderate recreational physical activity. Update 12/2013: set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.

If days of activity and time spend at different intensity levels are not missing, then
$$\text{GPAQ_rec_mod} = (\text{PAEA16} \times \text{PAEA17A_MIN})/7$$

Source variable(s):

In a typical week, ...

PAEA16. Days of moderate-intensity sports, fitness or recreational activities?

PAEA17A_MIN. Time spent doing moderate-intensity recreational activities (min)

11.2 GPAQ_rec_vig (Recreational physical (vigorous) activity (min/day))

This is a numeric variable that calculates the average amount of time spent per day doing a vigorous recreational physical activity. Update 12/2013: set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.

If days of activity and time spend at different intensity levels are not missing, then
$$\text{GPAQ_rec_vig} = (\text{PAEA12} \times \text{PAEA13A_MIN})/7$$

Source variable(s):

In a typical week, ...

PAEA12. Days of vigorous-intensity sports, fitness or recreational (leisure) activities?

PAEA13A_MIN. Time spent doing vigorous-intensity recreational activities (min)

11.3 GPAQ_rec (Recreational physical activity (min/day))

This is a numeric variable that finds the total time per day doing some form of recreational physical activity. Update 12/2013: set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.

If days of activity and time spend at different intensity levels are not missing, then
$$\text{GPAQ_REC} = \text{sum}(\text{GPAQ_REC_MOD}, \text{GPAQ_REC_VIG}).$$

Source variable(s):

GPAQ_rec_mod. Recreational physical (moderate) activity (min/day)

GPAQ_rec_vig. Recreational physical (vigorous) activity (min/day)

11.4 GPAQ_work_mod (Work-related (moderate) physical activity (min/day))

This is a numeric variable that calculates the average amount of time spent per day doing a moderate work-related physical activity. Update 12/2013: set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.

If days of activity and time spent at different intensity levels are not missing, then
$$\text{GPAQ_WORK_MOD} = (\text{PAEA5} \times \text{PAEA6A_MIN})/7$$

Source variable(s):

In a typical week, ...

PAEA5. Days of moderate-intensity activities as part of your work?

PAEA6A_MIN. Time spent doing moderate-intensity activities at work (min)

11.5 GPAQ_work_vig (Work-related (vigorous) physical activity (min/day))

This is a numeric variable that calculates the average amount of time spent per day doing a vigorous work-related physical activity. Update 12/2013: set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.

If days of activity and time spent at different intensity levels are not missing, then
$$\text{GPAQ_WORK_VIG} = (\text{PAEA2} \times \text{PAEA3A_MIN})/7$$

Source variable(s):

In a typical week, ...

PAEA2. Days of vigorous-intensity activities as part of your work?

PAEA3A_MIN. Time spent doing vigorous-intensity activities at work (min)

11.6 GPAQ_work (Work-related physical activity (min/day))

This is a numeric variable that finds the total time per day doing some form of work-related physical activity. Update 12/2013: set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.

If days of activity and time spent at different intensity levels are not missing, then
 $GPAQ_WORK = \text{sum}(GPAQ_WORK_MOD, GPAQ_WORK_VIG).$

Source variable(s):

GPAQ_work_mod. Work-related (moderate) physical activity (min/day)

GPAQ_work_vig. Work-related (vigorous) physical activity (min/day)

11.7 GPAQ_total_mod (Total physical (moderate) activity (min/day))

This is a numeric variable that finds the average amount of time spent per day doing moderate physical activity. Update 12/2013: set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.

If days of activity and time spent at different intensity levels are not missing, then
 $GPAQ_TOTAL_MOD = \text{sum}((PAEA5 \times PAEA6A_MIN), (PAEA8 \times PAEA9A_MIN), (PAEA16 \times PAEA17A_MIN))/7$

Source variable(s):

In a typical week, ...

PAEA5. Days of moderate-intensity activities as part of your work?

PAEA6A_MIN. Time spent doing moderate-intensity activities at work (min)

PAEA8. Days walk or bicycle to get to and from places?

PAEA9_MIN. Time spent walking or bicycling for travel (min)

PAEA16. Days of moderate-intensity sports, fitness or recreational activities?

PAEA17A_MIN. Time spent doing moderate-intensity recreational activities (min)

11.8 GPAQ_total_vig (Total physical (vigorous) activity (min/day))

This is a numeric variable that finds the average amount of time spent per day doing vigorous physical activity. Update 12/2013: set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.

If days of activity and time spent at different intensity levels are not missing, then
 $GPAQ_TOTAL_VIG = \text{sum}((PAEA2 \times PAEA3A_MIN), (PAEA12 \times PAEA13A_MIN))/7$

Source variable(s):

In a typical week, ...

PAEA2. Days of vigorous-intensity activities as part of your work?

PAEA3A_MIN. Time spent doing vigorous-intensity activities at work (min)

PAEA12. Days of vigorous-intensity sports, fitness or recreational (leisure) activities?

PAEA13A_MIN. Time spent doing vigorous-intensity recreational activities (min)

11.9 GPAQ_trsport (Transportation-related physical activity (min/day))

This is a numeric variable that calculates the average amount of time spent per day walking or biking to and from a place. Update 12/2013: set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.

If days of activity and time spend at different intensity levels are not missing, then

$$\text{GPAQ_TRSPORT} = (\text{PAEA8} \times \text{PAEA9A_MIN})/7$$

Source variable(s):

In a typical week, ...

PAEA8. Days walk or bicycle to get to and from places?

PAEA9_MIN. Time spent walking or bicycling for travel (min)

11.10 GPAQ_total (Total physical activity (min/day))

This is a numeric variable that shows the average amount of time doing some form of physical activity per day. Update 12/2013: set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.

If activity time for each domain is non-missing, then

$$\text{GPAQ_TOTAL} = \text{sum}(\text{GPAQ_WORK}, \text{GPAQ_TRSPORT}, \text{GPAQ_REC})$$

Source variable(s):

GPAQ_work. Work-related physical activity (min/day)

GPAQ_trsport. Transportation-related physical activity (min/day)

GPAQ_rec. Recreational physical activity (min/day)

11.11 GPAQ_total_MET (Total physical activity (MET-min/day))

This is a numeric variable that calculates the total amount of time spent doing some form of physical activity in a week. Update 12/2013: set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.

If activity time for each domain is non-missing, then

$$\begin{aligned} \text{GPAQ_TOTAL_MET} = & ((\text{PAEA2} \times \text{PAEA3A_MIN} \times 8) + (\text{PAEA5} \times \text{PAEA6A_MIN} \times 4) \\ & + (\text{PAEA8} \times \text{PAEA9A_MIN} \times 4) \\ & + (\text{PAEA12} \times \text{PAEA13A_MIN} \times 8) + (\text{PAEA16} \times \text{PAEA17A_MIN} \times 4))/7 \end{aligned}$$

Source variable(s):

PAEA2. Days per week of vigorous-intensity activities as part of your work

PAEA3A_MIN. Time (min) per day of vigorous-intensity activities at work

PAEA5. Days per week of moderate-intensity activities as part of your work

PAEA6A_MIN. Time (min) per day of moderate-intensity activities at work

PAEA8. Days per week of walk or bicycle to get to and from places

PAEA9A_MIN. Time (min) per day of walk or bicycle to get to and from places

PAEA12. Days per week of vigorous-intensity sports, fitness or recreational (leisure) activities

PAEA13A_MIN. Time (min) per day of vigorous-intensity sports, fitness or recreational (leisure) activities

PAEA16. Days per week of moderate-intensity sports, fitness or recreational (leisure) activities

PAEA17A_MIN. Time (min) per day of moderate-intensity sports, fitness or recreational (leisure) activities

11.12 GPAQ_MOD_WEEK (GPAQ - Total physical (moderate) activity (min/week))

This is a numeric variable that finds the average amount of time spent per week doing moderate physical activity.

$GPAQ_MOD_WEEK = GPAQ_TOTAL_MOD * 7$

Source variable(s):

GPAQ_TOTAL_MOD. Total physical (moderate) activity (min/day)

11.13 GPAQ_VIG_WEEK (GPAQ - Total physical (vigorous) activity (min/week))

This is a numeric variable that finds the average amount of time spent per week doing vigorous physical activity.

$GPAQ_VIG_WEEK = GPAQ_TOTAL_VIG * 7$

Source variable(s):

GPAQ_TOTAL_VIG. Total physical (vigorous) activity (min/day)

Source variable(s):

GPAQ_TOTAL_MOD. Total physical (moderate) activity (min/day)

GPAQ_TOTAL_VIG. Total physical (vigorous) activity (min/day)

11.15 GPAQ_level (Physical activity level)

This is a categorical variable which assigns a value of 1 assigned to those with high total physical activity levels, a value of 2 to those with moderate total physical activity levels and a value of 3 to those with low total physical activity levels. Update 12/2013: corrected a typo in code such that participants who were actually "moderate" were being set to "low". Set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.

Level of total physical activity	Physical activity cutoff value
High	<ul style="list-style-type: none"> • If vigorous PA (work+rec) \geq 3 days & GPAQ_total_MET \geq 1500 <li style="text-align: center;">OR • If moderate+vigorous PA \geq 7 days & GPAQ_total_MET \geq 3000
Moderate	<ul style="list-style-type: none"> • If vigorous PA (work+rec) \geq 3 days & GPAQ_total_vig \times 7 \geq 60 min <li style="text-align: center;">OR • If moderate PA (work+transport+rec) \geq 5 days & GPAQ_total_mod \times 7 \geq 150 min <li style="text-align: center;">OR • If moderate+vigorous PA \geq 5 days & GPAQ_total \times 7 \geq 600 min
Low	If the value does not reach the criteria for either high or moderate levels of physical activity

Source variable(s):

PAEA2. Days per week of vigorous-intensity activities as part of your work

PAEA5. Days per week of moderate-intensity activities as part of your work

PAEA8. Days per week of walk or bicycle to get to and from places

PAEA12. Days per week of vigorous-intensity sports, fitness or recreational (leisure) activities

PAEA16. Days per week of moderate-intensity sports, fitness or recreational (leisure) activities

GPAQ_TOTAL_VIG. Total physical (vigorous) activity (min/day)

GPAQ_TOTAL_MET. Total physical activity (MET-min/day)

GPAQ_TOTAL. Total physical (moderate) activity (min/day)

11.16 GPAQ_PAG2008 (Activity level per 2008 PA guidelines (categorical - 4 levels))

This is an ordinal variable that contains the level of physical activity per the 2008 PA guidelines (<http://www.health.gov/paguidelines/guidelines/default.aspx>). There are 4 mutually exclusive levels: high, medium, low or inactive based on all counts. Set participants to missing if they were missing GPAQ_MOD_WEEK and GPAQ_VIG_WEEK. Set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.

GPAQ_PAG2008	Definition from 2008 Physical Activity Guidelines (for adults)
4 = High activity	“more than the equivalent of 300 minutes of moderate-intensity physical activity a week” or more than 150 minutes of vigorous activity, or an equivalent combination of both. Activity should be performed in episodes of at least 10 minutes.
3 = Medium activity	“150 minutes to 300 (5 hours) minutes of moderate-intensity activity a week (or 75 to 150 minutes of vigorous-intensity physical activity a week)” or the equivalent combination of moderate and vigorous activity. Activity should be performed in episodes of at least 10 minutes.
2= Low activity	“activity beyond baseline but fewer than 150 minutes (2 hours and 30 minutes) of moderate-intensity physical activity a week or the equivalent amount (75 minutes, or 1 hour and 15 minutes) of vigorous-intensity activity” or the equivalent combination of moderate and vigorous activity.
1 = Inactive	“no activity beyond baseline activities of daily living”

Note that all borderline cases are given the higher level of activity. If none of these criteria are met, GPAQ_PAG2008 is set to missing.

```
if GPAQ_MOD_WEEK>=300 | GPAQ_VIG_WEEK>=150 |
GPAQ_MV_WEEK>=300 then GPAQ_PAG2008=4;

else if 150<=gpaq_mod_week<300 | 75<=gpaq_vig_week<150 |
150<=gpaq_mv_week<300 then GPAQ_PAG2008=3;

else if 1<=gpaq_mod_week<150 | 1<=gpaq_vig_week<75 |
1<=gpaq_mv_week<150 then GPAQ_PAG2008=2;

else if (0<=gpaq_mod_week<1 & 0<=gpaq_vig_week<1) then
GPAQ_PAG2008=1;

else GPAQ_PAG2008=.;
```

Response format: 1 = Inactive
2 = Low Activity
3 = Medium Activity
4 = High Activity

Variable GPAQ_MV_WEEK is an intermediate variable created for the purposes of this definition only. It is defined as below:

$GPAQ_MV_WEEK = (GPAQ_TOTAL_MOD * 7) + (2 * (GPAQ_TOTAL_VIG * 7))$

Source variable(s):

GPAQ_MOD_WEEK. GPAQ - Total physical (moderate) activity (min/week).

GPAQ_VIG_WEEK. GPAQ - Total physical (vigorous) activity (min/week).

GPAQ_MV_WEEK. GPAQ - Total physical activity (min/week).

11.17 GPAQ_PAG2008YN (Meets 2008 activity level guidelines (1=Yes, 0=No))

This is a 0/1 categorical variable that determines whether or not high or medium activity level has been met based on GPAQ_PAG2008 (all counts). Update 12/2013: changes reflected from GPAQ_PAG2008

$GPAQ_PAG2008YN = 1$ if $GPAQ_PAG2008 = 3$ or $GPAQ_PAG2008 = 4$
Else $GPAQ_PAG2008YN = 0$

Response format: 0 = No
1 = Yes

Source variable(s):

GPAQ_PAG2008. Activity level per 2008 PA guidelines (categorical - 4 levels)

12. CLINICAL CHARACTERISTICS

12.1 Hypertension (Hypertension (BP ≥ 140/90 and Med Use))

This is a 0/1 numeric variable. Hypertension is defined here if the systolic or diastolic BP is greater than or equal to 140/90 or if the participant is currently taking antihypertensive medications (scanned meds only). Participants without a blood pressure measurement and no medication use were assumed to be not hypertensive. This variable does NOT include self-report doctor's diagnosis of hypertension or self-report antihypertensive medication use.

5/2014 update: updates due to updated medication variables

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

Source variable(s):

SBPA5. Systolic blood pressure.

SBPA6. Diastolic blood pressure.

MED_ANTIHYPERT. Scanned medication was for treatment of hypertension.

12.2 Hypertension2 (Hypertension using NHANES definition)

This is a 0/1 numeric variable. Hypertension is defined here using the following NHANES definition: if the systolic or diastolic BP is greater than or equal to 140/90 or if the participant self-reported as currently taking antihypertensive medications. Participants without a blood pressure measurement and no medication use were assumed to be not hypertensive.

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

Source variable(s):

SBPA5. Systolic blood pressure.

SBPA6. Diastolic blood pressure.

MUEA33D. Were any of the medications you took during the last 4 weeks for high blood pressure or hypertension?

12.3 Hypertension_C4 (4-level grouped hypertension)

This is a numeric variable that classifies participants into 4 categories based on hypertension and hypertensive medication use (including both self-reported and Scanned medications): no hypertension, pre-hypertension, treated hypertension, and untreated hypertension.

HYPERTENSION2=0 and ((120≤SPBA5<140) or (80≤SPBA6<90)) → 2

Other HYPERTENSION2=0 → 1

HYPERTENSION2=1 and (MUEA33D=1 OR MED_ANTIHYPERT=1) → 3

Other HYPERTENSION2=1 → 4
Missing HYPERTENSION2→ missing

Response format: 1 = No hypertension
2 = Pre-hypertension
3 = Treated hypertension
4 = Untreated hypertension

Source variable(s):

Hypertension2. Hypertension using NHANES definition.
SBPA5. Systolic blood pressure.
SBPA6. Diastolic blood pressure.
MUEA33D. Self-reported medication for High blood pressure or hypertension.
Med_antihypert. Scanned Antihypertensive medication use.

12.4 Hypertension2_AHA (Hypertension using new ACC/AHA Guidelines definition)

This is a 0/1 numeric variable. Hypertension is defined here using new ACC/AHA Guidelines definition (New ACC/AHA Guideline reference:<https://www.acc.org/latest-in-cardiology/articles/2017/11/08/11/47/mon-5pm-bp-guideline-aha-2017>): if the systolic or diastolic BP is greater than or equal to 130/80 or if the participant self-reported as currently taking antihypertensive medications. Participants without a blood pressure measurement and no medication use were assumed to be not hypertensive.

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

Source variable(s):

SBPA5. Systolic blood pressure.
SBPA6. Diastolic blood pressure.
MUEA33D. Were any of the medications you took during the last 4 weeks for high blood pressure or hypertension?

12.5 Hypertension_AHA_C5 (5-level grouped hypertension using new ACC/AHA Guidelines definition)

This is a numeric variable that describes the stages of hypertension based on new ACC/AHA definition at 5 levels (None/Elevated/treated/untreated Stage 1/ untreated Stage 2). This definition takes into account systolic and diastolic blood pressure measured, and self-reported medications for high blood pressure or hypertension.

IF HYPERTENSION2_AHA=0 and ((120<=SBPA5<130) AND SBPA6<80) → 2
Otherwise HYPERTENSION2_AHA=0 → 1
HYPERTENSION2_AHA=1 and MUEA33D=1 → 3

Otherwise if HYPERTENSION2_AHA=1 and ((SBPA5>=140) OR (SBPA6>=90)) → 5
Otherwise if HYPERTENSION2_AHA=1 → 4
Missing HYPERTENSION2_AHA → missing

Response format: 1=Not hypertension
2=Elevated hypertension
3=Treated hypertension
4=Untreated stage 1 hypertension
5=Untreated stage 2 hypertension

Source variable(s):

Hypertension2. Hypertension using new ACC/AHA definition.

SBPA5. Systolic blood pressure.

SBPA6. Diastolic blood pressure.

MUEA33D. Self-reported medication for High blood pressure or hypertension.

12.6 Hypert_Awareness (Awareness of Hypertension - NHANES)

This is a 0/1 numeric variable and is defined by the self-report of a doctor's diagnosis of hypertension (MHEA1=1) for those who were defined as having hypertension using the NHANES definition (HYPERTENSION2=1) only. Note that participants who were missing MHEA1 were set to 0 (not aware).

Response format: 0=Not aware (No), or missing MHEA1
1=Aware (Yes)

Source variable(s):

MHEA1. Has a doctor ever said that you have high blood pressure or hypertension?

HYPERTENSION2. Hypertension using NHANES definition

12.7 Hypert_Treatment (Treatment of Hypertension - NHANES)

This is a 0/1 numeric variable and is defined by the participant's self-report awareness of hypertension (HYPERT_AWARENESS) and self-report use of antihypertensive medications (MUEA33D=1) for those who were defined as having hypertension using the NHANES definition (HYPERTENSION2=1) only. Note that participants who had missing data for hypertensive medication use were set to 0 (no medication use)

Response format: 0=No (or missing data for) antihypertensive med use (No)
1=Hypertensive med use (Yes)

Source variable(s):

MUEA33D. Were any of the medications you took during the last 4 weeks for high blood pressure or hypertension?

HYPERTENSION2. Hypertension using NHANES definition

HYPERT_AWARENESS. Awareness of Hypertension - NHANES

12.8 Hypert_Control (Control of Hypertension - NHANES)

This is a 0/1 numeric variable and is defined for participants whose blood pressure are below the cutoff of 140/90 (both systolic and diastolic must be less than the cutoff values) for those who were defined as having hypertension using the NHANES definition (HYPERTENSION2=1) only and also indicated whether/not they were receiving treatment for hypertension (HYPERT_TREATMENT= 0 or 1).

Response format: 0=Hypertension not under control (No)
1= Hypertension is under control (Yes)

Source variable(s):

SBPA5. Systolic blood pressure.

SBPA6. Diastolic blood pressure.

HYPERTENSION2. Hypertension using NHANES definition

HYPERT_TREATMENT. Treatment of Hypertension - NHANES

12.9 HOMA_IR (HOMA index of Insulin Resistance)

This is a numeric variable that defined the Homeostasis model assessment of insulin resistance as follows (LABA70 and INSULIN_FAST must be non-missing):

$HOMA_IR = LABA70 * INSULIN_FAST / 405.$

Source variable(s):

LABA70. Fasting Glucose in mg/dL

INSULIN_FAST. Insulin, fasting (converted to mU/L)

12.10 HOMA_B (HOMA index of Beta Cell Function)

This is a numeric variable that defined the Homeostasis model assessment of insulin resistance as follows:

If $LABA70 \leq 63$ then set HOMA_B as missing, otherwise:

$HOMA_B = 360 * INSULIN_FAST / (LABA70 - 63).$

Source variable(s):

LABA70. Fasting Glucose in mg/dL

INSULIN_FAST. Insulin, fasting (converted to mU/L)

12.11 Diabetes_self (Diabetes – self-report only)

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

If the MHEA form is missing then Diabetes_self =.

If mhea16=1 and mhea16b=(0 or Missing) then Diabetes_self =1.

Else if mhea16 is missing then Diabetes_Self= missing.

Otherwise Diabetes_Self=0

5/2014 update: revised to code missing MHEA16 as no and MHEA16=.Q as missing instead of zero

Response format: 0 = No
1 = Yes

Source variable(s):

MHEA16. Has a doctor ever said that you have diabetes.
MHEA16b. Was this during pregnancy.

12.12 Diabetes_Lab (3-level grouped Diabetes – Lab ADA Guideline)

This is a numeric variable that describes the stages of diabetes. This definition takes into account serum glucose levels adjusted for fasting time and, if available, post-OGTT glucose levels, and A1C percentages. This variable does NOT include any medication use variables.

This variable is missing if fasting glucose, post-OGTT, and A1C are missing.

If (fasting time > 8 AND fasting glucose \geq 126 mg/dL) or (fasting time \leq 8 AND fasting glucose \geq 200 mg/dL) or (post-OGTT glucose \geq 200 mg/dL) or (A1C \geq 6.5%) or then DIABETES_Lab=3

Else if (fasting time > 8 AND fasting glucose in range 100-125 mg/dL) or (post-OGTT glucose in range 140-199 mg/dL) or (5.7% \leq A1C<6.5%) then DIABETES_Lab=2
Else DIABETES_Lab=1

Response format: 1 = normal glucose regulation
2 = impaired glucose tolerance
3 = diabetes

Source variable(s):

LABA70. Fasting glucose (mg/dL)
LABA71. Glucose, post OGGT (mg/dL)
LABA72. % Glycosylated Hemoglobin (A1C)
FASTTIME. Elapsed time (hrs) between the time the participant last consumed anything and the blood draw

12.13 SelfMed_Antidiab (Anti-Diabetics Medication used – Self-reported)

This is a 0/1 numeric variable and is defined by the participant's self-report use of anti-diabetics medications (Yes if MUEA33C=1, No if MUEA33C=0, after reset missing to zero when confirmed not taking any medications, MUEA2=1). Note that participants who had missing data or unsure (MUEA33C=9) for antidiabetics medication use were set to missing.

Response format: 0=No antidiabetics med reported (No)
1=Self-reported use antidiabetics med (Yes)

Source variable(s):

MUEA33C. Were any of the medications you took during the last 4 weeks for high blood sugar or diabetes?

MUEA2. Not bringing any of the medications you took during the last 4 weeks, because you forgot/unable to, or took no medication?

12.14 Diabetes1 (3-level grouped Diabetes ver.1)

This is a numeric variable that describes the stages of diabetes. This definition takes into account serum glucose levels adjusted for fasting time and, if available, post-OGTT glucose levels and scanned/transcribed anti-diabetic medication use.

This variable is missing if fasting glucose, post-OGTT and Med_Antidiab are missing. Note that participants with no glucose lab data and no anti-diabetic medication use were assumed to have normal glucose regulation.

If (fasting time > 8 AND fasting glucose \geq 126) or (fasting time \leq 8 AND fasting glucose \geq 200) or (post-OGTT glucose \geq 200) or (Med_Antidiab=1) then DIABETES1=3
Else if (fasting time > 8 AND fasting glucose in range 100-125) or (post-OGTT glucose in range 140-199) then DIABETES1=2
Else DIABETES1=1

5/2014 update: updates due to updated medication variables

Response format: 1 = normal glucose regulation
2 = impaired glucose tolerance
3 = diabetes

Source variable(s):

LABA70. Fasting glucose (mg/dL)

LABA71. Glucose, post OGGT (mg/dL)

FASTTIME. Elapsed time (hrs) between the time the participant last consumed anything and the blood draw

Med_Antidiab. Scanned/transcribed anti-diabetic medication use

12.15 Diabetes2 (3-level grouped Diabetes – ADA Guideline plus scanned Med)

This is a numeric variable that describes the stages of diabetes based on American Diabetes Association definition (*Diabetes Care* 2010;33:S62-69). This definition takes into account serum glucose levels adjusted for fasting time and, if available, post-OGTT glucose levels, A1C percentages, and scanned/transcribed anti-diabetic medication use.

This variable is missing if fasting glucose, post-OGTT, A1C, and Med_Antidiab are missing. Note that participants with no glucose lab data and no anti-diabetic medication use were assumed to have normal glucose regulation.

If (fasting time > 8 AND fasting glucose \geq 126 mg/dL) or (fasting time \leq 8 AND fasting glucose \geq 200 mg/dL) or (post-OGTT glucose \geq 200 mg/dL) or (A1C \geq 6.5%) or (Med_Antidiab=1) then DIABETES2=3

Else if (fasting time > 8 AND fasting glucose in range 100-125 mg/dL) or (post-OGTT glucose in range 140-199 mg/dL) or (5.7%≤A1C<6.5%) then DIABETES2=2
Else DIABETES2=1

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

Source variable(s):

LABA70. Fasting glucose (mg/dL)

LABA71. Glucose, post OGGT (mg/dL)

LABA72. % Glycosylated Hemoglobin (A1C)

FASTTIME. Elapsed time (hrs) between the time the participant last consumed anything and the blood draw

Med_Antidiab. Scanned/transcribed anti-diabetic medication use

12.16 Diabetes2_Indicator (Diabetes 2 Indicator (ADA))

This is a 0/1 indicator variable based on the Diabetes variable (version 2) stages of diabetes based on American Diabetes Association definition (*Diabetes Care* 2010;33:S62-69).

Missing, if Diabetes2 is missing;

1, if Diabetes2=3;

0, otherwise.

Response format: 0 = Non-diabetic (No)
1 = Diabetic (Yes)

Source variable(s):

DIABETES2. Diabetes2 (3-level grouped Diabetes – ADA)

12.17 Diabetes3 (3-level grouped Diabetes – ADA Guideline plus Self-reported Diabetes)

This is a numeric variable that describes the stages of diabetes based on American Diabetes Association definition (*Diabetes Care* 2010;33:S62-69) with an additional self-report criterion. This definition takes into account serum glucose levels adjusted for fasting time and, if available, post-OGTT glucose levels, A1C percentages, and self-report of diabetes

If Diabetes_Self=1 or Diabetes_Lab=3 then Diabetes3=3.

Else if Diabetes_Lab=2 then Diabetes3=2.

Else if Diabetes_Self=missing and Diabetes_Lab=(1 or missing) then Diabetes3=missing.

Otherwise Diabetes3=1.

5/2014 update: updates due to updated DIABETES_SELF

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

Source variable(s):

Diabetes_Lab. 3-level grouped Diabetes – Lab criterion only
Diabetes_Self. Diabetes – self-report only

12.18 Diabetes_C4 (4-level grouped diabetes (none/pre/treated/untreated))

This 4 category variable assesses diabetes using DIABETES3 and medication use based on MUEA33C (self-reported medication for high blood sugar or diabetes) and MED_ANTIDIAB (scanned antidiabetic medication)

DIABETES3=1 → 1
DIABETES3=2 → 2
DIABETES3=3 and (MUEA33C=1 or MED_ANTIDIAB=1) → 3
Other DIABETES3=3 → 4
Missing DIABETES3 → missing

5/2014 update: updates due to updated DIABETES3 and medication variables

Response format: 1=No diabetes
2=Pre-diabetes
3=Treated diabetes
4=Untreated diabetes

Source variable(s):

DIABETES3. 3-level grouped Diabetes - includes self-report
MUEA33C. Self-reported medication for high blood sugar or diabetes
MED_ANTIDIAB. Scanned antidiabetics

12.19 Diab_Diag (Diabetes Diagnosis)

This is a categorical variable that describes the diagnoses of diabetes and whether it was diagnosed before baseline, not diagnosed before baseline but possibly detected from lab data, or as non-diabetic. Note that participants who were diabetic but had missing data for self-reported diabetes were assumed to be undiagnosed diabetics.

If Diabetes_Self=1 then Diab_diag=1.
Else if Diabetes_Self not =1 and Diabetes3=3 then Diab_diag=2.
Else if Diabetes_Self=missing and Diabetes3= missing then Diab_diag=missing.
Otherwise Diab_diag=3

5/2014 update: updates due to updated DIABETES_SELF and DIABETES3

Response format: 1 = Diagnosed diabetic before baseline

2 = Undiagnosed diabetic (or missing self-report data)
3 = Non-diabetic

Source variable(s):

Diabetes3. 3-level grouped Diabetes – includes self-report
Diabetes_Self. Diabetes – self-report only

12.20 Diab_FamHist (Family History of Diabetes)

This is a 0/1 numeric variable that describes the family history of diabetes.

If (MHEA16e and MHEA16f and MHEA16g) all missing then Diab_FamHist= missing.
Else If MHEA16e=1 or MHEA16f=1 or MHEA16g=1 then Diab_Fam_Hist=1;
Otherwise Diab_Fam_Hist=0.

Response format: 0=No family history of diabetes (No)
1=Family history of diabetes (Yes)

Source variable(s):

MHEA16e. Mother had diabetes
MHEA16f. Father had diabetes
MHEA16g. Brother/sister had diabetes

12.21 DM_Aware (Diabetes Awareness at Baseline)

This is a 0/1 numeric variable that describes the awareness of diabetes. Defined only for self-reported diabetics.

if DIABETES_SELF missing AND (MUEA33c missing or =9) then DM_Aware=.;
else if DIABETES_SELF=1 OR muea33c=1 then DM_Aware=1;
else DM_Aware=0;

Set to missing for nondiabetics.

Response format: 0=Not aware of diabetes
1=Aware of diabetes

Source variable(s):

MUEA33c. High blood sugar or diabetes – self-reported med use
DIABETES_SELF. Diabetes – self-report only
DIABETES3. Diabetes ADA with self-reported DM

12.22 DM_Control (Diabetes controlled: DM Classified at Baseline and A1C < 7%)

This is a 0/1 numeric variable that indicates controlled diabetes. Defined only for diabetics.

if DIABETES3^=3 then DM_CONTROL=.;
else if DIABETES3=3 & .Z<lab72<7 then DM_CONTROL=1;

else DM_CONTROL=0;

5/2014 update: set to missing for nondiabetics

Response format: 0=Uncontrolled diabetes
 1=Controlled diabetes

Source variable(s):

Diabetes3. 3-level grouped Diabetes – includes self-report
LABA72. A1C

12.23 Diabetes4 (3-level grouped Diabetes – ADA Guideline plus self-reported Med)

This is a numeric variable that describes the stages of diabetes based on American Diabetes Association definition (*Diabetes Care* 2010;33:S62-69). This definition uses same definition as Diabetes2, except using self-reported medication instead of scanned medication.

If selfmed_antidiab=1 | diabetes_lab=3 then DIABETES4=3;
Else if diabetes_lab=2 then DIABETES4=2;
Else if selfmed_antidiab=. & diabetes_lab=. then DIABETES4=.;
Else DIABETES4=1;

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

Source variable(s):

Diabetes_lab. ADA Guideline Lab criteria based on FPG, OGTT and A1c
SelfMed_Antidiab. Self-reported use antidiabetics medication.

12.24 Diabetes5 (3-level grouped Diabetes – ADA Guideline plus Self-reported diabetes diagnosis or Self-reported Med)

This is a numeric variable that describes the stages of diabetes based on ADA Guideline lab criteria, Self-reported diabetes diagnosis or Self-reported antidiabetics med.

If diabetes_self=1 | selfmed_antidiab=1 | diabetes_lab=3 then DIABETES5=3;
Else if diabetes_lab=2 then DIABETES5=2;
Else if diabetes_self=. & selfmed_antidiab=. & diabetes_lab=. then DIABETES5=.;
Else DIABETES5=1;

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

Source variable(s):

Diabetes_lab. ADA Guideline Lab criteria based on FPG, OGTT and A1c
Diabetes_Self. Self-reported diabetes diagnosis.
SelfMed_Antidiab. Self-reported use antidiabetics medication.

12.25 PreCHD_No_Angina (Prevalent Cardiovascular Heart Disease)

This is a 0/1 variable that combines ECG reports of possible old MI as well as self-report of heart attack, or procedure (angioplasty, stent, bypass).

```
if ecga267^=' ' | ecga268^=' ' | mhea4=1 | mhea9=1 then PreCHD_No_Angina=1;  
else if ecga267=' ' & ecga268=' ' & mhea4 in(.,.Q) & mhea9 in(.,.Q) then  
PreCHD_No_Angina=.;  
else PreCHD_No_Angina=0;
```

Response format: 0=No Prevalent CHD (No)
1=Prevalent CHD (Yes)

Source variable(s):

ECGA267. Major Q Wave abnormalities [old Myocardial Infarction (MI)]

ECGA268. Minor Q, QS waves with ST, T abnormalities (possible old MI)

MHEA4. Has a doctor ever said that you had a heart attack?

MHEA9. Have you had a balloon angioplasty, a stent, or bypass surgery to the arteries in your heart to improve the blood flow to your heart?

12.26 PreCHD_Angina (Prevalent Cardiovascular Heart Disease, including angina)

This is a 0/1 variable that combines ECG reports of possible old MI as well as self-report of angina, heart attack, or procedure (angioplasty, stent, bypass).

```
if ecga267^=' ' | ecga268^=' ' | mhea3=1 | mhea4=1 | mhea9=1 then PreCHD_Angina=1;  
else if ecga267=' ' & ecga268=' ' & mhea3 in(.,.Q) & mhea4 in(.,.Q) & mhea9 in(.,.Q) then  
PreCHD_Angina =.;  
else PreCHD_Angina =0;
```

Response format: 0=No Prevalent CHD (No)
1=Prevalent CHD (Yes)

Source variable(s):

ECGA267. Major Q Wave abnormalities [old Myocardial Infarction (MI)]

ECGA268. Minor Q, QS waves with ST, T abnormalities (possible old MI)

MHEA3. Has a doctor ever said that you have angina?

MHEA4. Has a doctor ever said that you had a heart attack?

MHEA9. Have you had a balloon angioplasty, a stent, or bypass surgery to the arteries in your heart to improve the blood flow to your heart?

12.27 Ever_Angina_Relative (At least one relative had report of angina)

This is a 0/1 variable that combines self-report of angina for immediate family relatives, i.e. mother, father, and sibling(s).

If MHEA3 is non-missing then:
IF MHEA3b, MHEA3c, and MHEA3d are all missing then Ever_angina_relative = missing.

IF MHEA3b=1 or MHEA3c=1 or MHEA3d=1 then Ever_angina_relative = 1
Otherwise Ever_angina_relative = 0.

Response format: 0=No family history of angina (No)
1=Family history of angina (Yes)

Source variable(s):

MHEA3. Has a doctor ever said that you have angina?

MHEA3b. Has a doctor ever said that these relatives have angina: Mother

MHEA3c. Has a doctor ever said that these relatives have angina: Father

MHEA3d. Has a doctor ever said that these relatives have angina: Brother(s) or Sister(s)

12.28 Ever_MI_Relative (At least one relative had report of Heart Attack)

This is a 0/1 variable that combines self-report of heart attack for immediate family relatives, i.e. mother, father, and sibling(s).

If MHEA4 is non-missing then:

IF MHEA4B1, MHEA4C1, and MHEA4D1 are all missing then Ever_MI_relative = missing.

IF MHEA4B1=1 or MHEA4C1=1 or MHEA4D1=1 then Ever_MI_relative = 1

Otherwise Ever_MI_relative = 0.

Response format: 0=No family history of heart attack (No)
1=Family history of heart attack (Yes)

Source variable(s):

MHEA4. Has a doctor ever said that you have had a heart attack?

MHEA4B1. Has a doctor ever said that these relatives have had a heart attack: Mother

MHEA4C1. Has a doctor ever said that these relatives have had a heart attack: Father

MHEA4D1. Has a doctor ever said that these relatives have had a heart attack: Brother(s) or Sister(s)

12.29 Ever_CABG_Relative (At least one relative had report of CABG or angioplasty)

This is a 0/1 variable that combines self-report of CABG or angioplasty for immediate family relatives, i.e. mother, father, and sibling(s).

If MHEA9 is non-missing then:

IF MHEA9a, MHEA9b, and MHEA9c are all missing then Ever_CABG_relative = missing.

IF MHEA9a=1 or MHEA9b=1 or MHEA9c=1 then Ever_CABG_relative = 1

Otherwise Ever_angina_relative = 0.

Response format: 0=No family history of CABG or angioplasty (No)

1=Family history of CABG or angioplasty (Yes)

Source variable(s):

MHEA9. Have you had a balloon angioplasty, a stent, or bypass surgery to the arteries in your heart to improve the blood flow to your heart?

MHEA9a. Has a doctor ever said that these relatives have CABG or angioplasty: Mother

MHEA9b. Has a doctor ever said that these relatives have CABG or angioplasty: Father

MHEA9c. Has a doctor ever said that these relatives have CABG or angioplasty:

Brother(s) or Sister(s)

12.30 FH_CHD (Family history of CHD including self-reported angina)

This is a 0/1 variable that combines the self-reports of angina, heart attack, or procedure (angioplasty, stent, bypass) for immediate family relatives, i.e. mother, father, and sibling(s).

IF Ever_angina_relative, Ever_MI_relative, and Ever_CABG_relative are all missing then FH_CHD = missing.

Else if Ever_angina_relative=1 or Ever_MI_relative=1 or Ever_CABG_relative=1 then FH_CHD = 1.

Otherwise FH_CHD = 0

Response format: 0=No family history of CHD (No)

1=Family history of CHD (Yes)

Source variable(s):

Ever_Angina_relative. At least one relative had report of angina

Ever_MI_relative. At least one relative had report of Heart Attack

Ever_CABG_relative. At least one relative had report of CABG or angioplasty

12.31 CHD_SELF (Self-reported CHD)

This is a 0/1 variable that combines the self-reports of coronary revascularization or heart attack.

If mhea4=1 | mhea9=1 then CHD_SELF=1;

Else if mhea4 in(.,.Q) & mhea9 in(.,.Q) then CHD_SELF=.;

Else CHD_SELF=0;

Response format: 0=No CHD (No)

1= CHD (Yes)

Source variable(s):

MHEA4. Has a doctor ever said that you have had a heart attack?

MHEA9. Have you ever had a CABG, stent, or balloon angioplasty?

12.32 CDCR (Cerebrovascular Disease or Carotid Revascularization, self-report)

This is a 0/1 variable that checks for medical history of stroke, mini-stroke or TIA, or balloon angioplasty or surgery to the arteries in the neck, as follows:

If MHEA10=1 | MHEA11=1 | MHEA12=1 then CDCR=1;
Else if MHEA10≤.Z & MHEA11≤.Z & MHEA12≤.Z then CDCR=.;
Else CDCR=0;

Response format: 0 = No CDCR (No)
1 = CDCR (Yes)

Source variable(s):

MHEA10. Has a doctor ever said that you had a stroke?

MHEA11. Has a doctor ever said that you had a mini-stroke or TIA (transient ischemic attack)?

MHEA12. Have you had a balloon angioplasty or surgery to the arteries of your neck to prevent or correct a stroke?

12.33 METS_NCEP (Metabolic Syndrome – NCEP-ATP III)

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

The Adult Treat 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 based on the updated 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:

Elevated BP plusMeds: SBPA5 ≥ 130 or SBPA6 ≥ 85 mm Hg or use of anti-hypertension medications (Med_Antihypert=1) (note: requires non-missing SPB, DBP, or medication use).

High Trig: LABA67 ≥150 mg/dL

Low HDL: LABA68 < 40 mg/dl (men), LABA68 < 50 mg/dl (women)

IFG_NCEP: LABA70 ≥ 100 mg/dl or use of anti-diabetic medications (MED_ANTIDIAB=1)

Abdominal obesity_NCEP: ANTA10A ≥102 cm (men), ANTA10A ≥88 cm (women)

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

If 3 or more components are missing components, then METS_NCEP should be missing.

Otherwise, METS_NCEP = 0 if sum of five components is 0, 1, 2

= 1 if sum of five components is 3, 4, 5

5/2014 update: revised to use GENDER instead of PIEA1

Source variable(s):

GENDER Gender

ANTA10A. Girth - waist
LABA67. Triglycerides (mg/dL)
LABA68. HDL cholesterol (mg/dL)
LABA70. Fasting glucose (mg/dL)
SBPA5. Systolic blood pressure
SBPA6. Diastolic blood pressure
MED_ANTIHYPERT. Anti-hypertension medication use
MED_ANTIDIAB. Anti-diabetic medication use

12.34 METS_NCEP2 (Metabolic Syndrome w/ self-reported meds – NCEP-ATP III)

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

The diagnosis of metabolic syndrome is made when three or more of the following risk factors are present (*indicates variable substitutions from original definition):

*Elevated BP SelfMeds: SBPA5 \geq 130 or SBPA6 \geq 85 mm Hg or self-report use of anti-hypertension medications (MUEA33D=1) (note: requires non-missing SPB, DBP, or self-report medication use).

High Trig: LABA67 \geq 150 mg/dL

Low HDL: LABA68 < 40 mg/dl (men), LABA68 < 50 mg/dl (women)

*IFG NCEP SelfMeds: LABA70 \geq 100 mg/dl or self-report use of anti-diabetic medications (MUEA33C=1)

Abdominal obesity NCEP: ANTA10A \geq 102 cm (men), ANTA10A \geq 88 cm (women)

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

If 3 or more components are missing components, then METS_NCEP2 should be missing.

Otherwise, METS_NCEP2 = 0 if sum of five components is 0, 1, 2
= 1 if sum of five components is 3, 4, 5

12/2019 update: revised to use GENDER instead of PIEA1

Source variable(s):

GENDER. Gender

ANTA10A. Girth - waist

LABA67. Triglycerides (mg/dL)

LABA68. HDL cholesterol (mg/dL)

LABA70. Fasting glucose (mg/dL)

SBPA5. Systolic blood pressure

SBPA6. Diastolic blood pressure

MUEA33C. Self-report anti-diabetic medication use

MUEA33D. Self-report anti-hypertensive medication use

12.35 METS_IDF (Metabolic Syndrome – IDF)

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. The IDF definition emphasizes central obesity because it is “independently” associated with each of the other metabolic syndrome components, as well as insulin resistance. Another difference between the IDF and NCEP-ATP III definitions is the recommended use of ethnic-specific waist girth thresholds (thresholds for Asian populations are lower than those for European and North American groups). Further, the ATP III definition includes 110 mg/dL cutoff for fasting plasma glucose (although the American Diabetes Association recently decided to lower the threshold for impaired fasting glucose to 100 mg/dL).

The IDF diagnosis of metabolic syndrome is made:

- a) Abdominal obesity IDF: ANTA10A \geq 94 (men), ANTA10A \geq 80 cm (women)
- b) Plus any 2 of the following:
 - High Trig plusMeds: LABA67 \geq 150 mg/dL or treatment for this lipid abnormality (Med_ Fibares_Nicoacid =1)
 - Low HDL plusMeds: LABA68 $<$ 40 mg/dL (men), LABA69 $<$ 50 mg/dL (women) or specified treatment for this lipid abnormality (Med_ Fibares_Nicoacid =1)
 - Elevated BP plusMeds: SBPA5 \geq 130 or SBPA6 \geq 85 mm Hg or use of anti-hypertension medications (Med_Antihypert =1) (note: requires non-missing SPB, DBP, or medication use).
 - IFG IDF: LABA70 \geq 100 mg/dl, or previous diagnosis of diabetes (MHEA16=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.

5/2014 update: corrected code to require abdominal obesity, revised to use GENDER instead of PIEA1

Source variable(s):

GENDER. Gender

MHEA16. Has a doctor ever said that you have diabetes (high sugar in blood or urine)

ANTA10A. Girth – waist

LABA67. Triglycerides (mg/dL)

LABA68. HDL cholesterol (mg/dL)

LABA70. Fasting glucose (mg/dL)

SBPA5. Systolic blood pressure
SBPA6. Diastolic blood pressure
MED_Fibares_Nicoacid. Fibrin/Nicotinic Acids (trt of TG and HDL)
MED_ANTIDIAB. Anti-diabetic medication use
MED_ANTIHYPERT. Anti-hypertensive medication use

12.36 MI_ECG (MI defined by ECG results)

This is a 0/1 variable that checks some ECG measures to detect possible old MIs. Participants without ECGs are set to missing.

```
else if ecga267 ne '' | ecga268 ne '' then MI_ECG=1;  
else if ecga267 = '' & ecga268 = '' then MI_ECG=0;
```

Source variable(s):

ECGA267. Major Q Wave abnormalities [old Myocardial Infarction (MI)]
ECGA268. Minor Q, QS waves with ST, T abnormalities (possible old MI)

12.37 Stroke (Prevalent Stroke, self-report)

This is a 0/1 variable that checks for a self-report medical history of stroke as follows:
STROKE=MHEA10;

Response format: 0=No Prevalent Stroke (No)
1= Prevalent Stroke (Yes)

Source variable(s):

MHEA10. Has a doctor ever said that you had a stroke?

12.38 Stroke_TIA (Prevalent Stroke or TIA, self-report)

This is a 0/1 variable that checks for medical history of stroke or mini-stroke or TIA as follows:

```
if MHEA10=1 | MHEA11=1 then STROKE_TIA=1;  
else if MHEA10<=.Z & MHEA11<=.Z then STROKE_TIA=.;  
else STROKE_TIA=0;
```

Response format: 0 = No Prevalent Stroke/TIA (No)
1 = Prevalent Stroke/TIA (Yes)

Source variable(s):

MHEA10. Has a doctor ever said that you had a stroke?
MHEA11. Has a doctor ever said that you had a mini-stroke or TIA (transient ischemic attack)?

12.39 FH_Stroke (Family History of Stroke)

This is a 0/1 variable that combines self-report of stroke for immediate family relatives, i.e. mother, father, and sibling(s).

If MHEA10 is non-missing then:

IF MHEA10a, MHEA10b, and MHEA10c are all missing then FH_Stroke = missing.

IF MHEA10a=1 or MHEA10b=1 or MHEA10c=1 then FH_Stroke = 1

Otherwise FH_Stroke = 0.

Response format: 0=No family history of stroke (No)
1=Family history of stroke (Yes)

Source variable(s):

MHEA10. Has a doctor ever said that you had a stroke?

MHEA10a. Has a doctor ever said that these relatives had a stroke: Mother

MHEA10b. Has a doctor ever said that these relatives had a stroke: Father

MHEA10c. Has a doctor ever said that these relatives had a stroke: Brother(s) or Sister(s)

12.40 PRECVD (Prevalent CVD (CHD or Stroke, not counting Angina or TIA))

This is a 0/1 variable that indicates prevalent cardiovascular disease, including CHD and stroke, without counting angina or TIA. as follows:

If PRECHD_NO_ANGINA<=.Z & STROKE<=.Z then PRECVD=.;
Else if PRECHD_NO_ANGINA=1 | STROKE=1 then PRECVD=1;
Else PRECVD=0;

Response format: 0=No Prevalent CVD (No)
1= Prevalent CVD (Yes)

Source variable(s):

PRECHD_NO_ANGINA. Prevalent CHD, not counting Angina.

STROKE. Prevalent Stroke, self-report.

12.41 Cigarette_Use (1=Never,2=Former,3=Current)

Current, former, and never cigarette use.

Response format: 1 = "Never"
2 = "Former"
3 = "Current"

if TBEA1=1 & TBEA3 in(1,2) then CIGARETTE_USE=3;
else if TBEA1=1 & TBEA3 in(3) then CIGARETTE_USE=2;
else if TBEA1=0 then CIGARETTE_USE=1;

Source variable(s):

TBEA1. Have you ever smoked at least 100 cigarettes in your entire life?

TBEA3. Do you now smoke daily, some days or not at all?

12.42 Current_Smoker (Current Smoker (from CIGARETTE_USE))

This is a 0/1 indicator variable to denote current smoker (CIGARETTE_USE=3) status using the variable CIGARETTE_USE.

Response format: 0 = No (i.e., never or former smoker)
1 = Yes (i.e., current smoker)

Source variable(s):

CIGARETTE_USE. Cigarette_Use (1=Never,2=Former,3=Current)

12.43 Exposure_Year (Adjusted Years Smoked)

This is a numeric variable that calculates the years a participant smoked cigarettes. This is calculated using the difference from the age the participant began smoking to either:

- a. the participant's age at baseline examination or
- b. the age that the participant quit smoking (if applicable)

This variable is also adjusted for any years that the participant quit smoking during the above calculated smoking time period (Note: TBEA4B, 5C, and 6B are = 0 if the participant did not quit smoking during this time period). If the final calculation is between 0 and 1, then the variable is set to =1. If the number of years quit was missing, it was assumed to be zero.

Five cases for computing years of cigarette smoking exposure:

- 1) Current smoker, Daily: If (TBEA1=1 and TBEA3=1) then Exposure_year = (Age – TBEA2 – TBEA4B)
- 2) Current smoker, Some days: If (TBEA1=1 and TBEA3=2) then Exposure_year = (Age – TBEA2 – TBEA5C)
- 3) Former smoker: If (TBEA1=1 and TBEA3=3) then Exposure_Year = (TBEA6 – TBEA2 – TBEA6B)
- 4) Never Smoked: If (TBEA1=0) then Exposure_year = 0
- 5) All Other cases: Exposure_year = .

5/2014 update: revised to treat TBEA2=99 or 1 as missing. Corrected code to set never smokers to zero, former and current smokers with negative values to missing and to subtract TBEA2 for former smokers

Source variable(s):

- TBEA1. Have you ever smoked at least 100 cigarettes in your entire life?
TBEA2. How old were you when you first started to smoke cigarettes fairly regularly?
TBEA3. Do you now smoke daily, some days or not at all?
TBEA4A. Did you ever quit smoking for 6 months or longer? (Daily)
TBEA4B. For how many years in total did you quit smoking? (Daily)
TBEA5B. Did you ever quit smoking for 6 months or longer? (Some days)
TBEA5C. For how many years in total did you quit smoking? (Some days)
TBEA6. How old were you when you completely stopped smoking?

TBEA6A. When you were a smoker, did you ever quit smoking for 6 months or longer before you

TBEA6B. During the time that you were a smoker, for how many years in total did you quit smoking? (Former)

Age. Age in years at the time of participant's clinic visit

12.44 Cigarettes_Year (Adjusted Cigarette Years)

This is a numeric variable that calculates the cigarette years as the number of exposure years multiplied by the average number of cigarettes smoked per day. Set to missing for participants who said yes to TBEA1 and 0 for TBEA9..

5/2014 update: updates due to changes in Exposure_Year, set current smokers with TBEA9=0 to missing

Source variable(s):

TBEA1. Have you ever smoked at least 100 cigarettes in your entire life?

TBEA9. Of the entire time you have or had smoked, on average how many cigarettes do you or did you smoke per day?

Exposure_Year. Adjusted Years Smoked

12.45 Cigarette_Pack_Years (Cigarette Pack Years)

This is a numeric variable that calculates the cigarette pack years as the number of exposure years multiplied by the average number of cigarettes smoked per day (Cigarettes_Year), divided by 20.

$$\text{Cigarette_Pack_years} = \text{Cigarettes_Year}/20$$

5/2014 update: updates due to changes in Exposure_Year

Source variable(s):

Cigarettes_Year. Adjusted Cigarette Years

12.46 Alcohol_Use (1=Never,2=Former,3=Current)

Current, former, and never alcohol use.

Response format: 1 = "Never"
2 = "Former"
3 = "Current"

if ALEA1=1 then ALCOHOL_USE=3;
else if ALEA1=0 & ALEA7=1 then ALCOHOL_USE=2;
else if ALEA1=0 & ALEA7=0 then ALCOHOL_USE=1;

Source variable(s):

ALEA1. Do you presently drink alcoholic beverages?

ALEA7. Did you ever drink alcohol?

12.47 TotalDrinks_Per_Week (Total Weekly Alcohol Consumption)

This is a numeric variable that determines the total alcohol Intake of the numbers of drinks per week only for current drinkers.

```
if ALEA1^=1 then ALCOHOL_INTAKE=.;  
else ALCOHOL_INTAKE=sum(of ALEA2-ALEA5);
```

Source variable(s):

ALEA1. Do you presently drink alcoholic beverages?

ALEA2. How many glasses of red wine do you usually have per week?

ALEA3. How many glasses of white wine do you usually have per week?

ALEA4. How many cans, bottles, or glasses of beer do you usually have per week?

ALEA5. How many drinks of liquor, spirits, or mixed drinks do you usually have per week?

12.48 Alcohol_Use_Disorder (Alcohol Use Disorder Risk)

Based on gender-specific cutoffs for weekly alcohol use (see below) provided by the National Institute for Alcohol Abuse and Alcoholism (www.niaaa.nih.gov), ALCOHOL_USE_DISORDER provides a risk level for developing an alcohol use disorder (craving, loss of control, physical dependence, tolerance, etc.). Note that the NIAAA criteria includes binge drinking (i.e. drinks on any single day) levels that were not collected in HCHS/SOL. ALCOHOL_USE_DISORDER sets former drinkers to missing with the reasoning that they may meet criteria for an alcohol use disorder, but the assessment of that risk is not possible based on current use.

Gender	Weekly alcohol use	Alcohol_Use_Disorder
F	never used alcohol	Non-drinker
F	current use <7 drinks/week	Low risk drinker
F	7+ drinks/week	At-risk drinker
M	never used alcohol	Non-drinker
M	current use <14 drinks/week	Low risk drinker
M	14+ drinks/week	At-risk drinker

Response format: 0 = Non-drinker
1 = Low risk drinker
2 = At-risk drinker

Source variable(s):

Gender. Participant Gender

Alcohol_Use. (1=Never,2=Former,3=Current)

TotalDrinks_Per_Week. Total Weekly Alcohol Consumption

12.49 Alcohol_Use_Level (Alcohol Use Disorder Risk - NIAAA)

Using the weekly alcohol use cutoffs from ALCOHOL_USE_DISORDER, ALCOHOL_USE_LEVEL combines never users former users into a group called "no current use". Definitions of "low level use" and "high level use" correspond to those of "low risk drinker" and "at-risk drinker" from ALCOHOL_USE_DISORDER

Gender	Weekly alcohol use	Alcohol_Use_Level
F	no current use (never or former users)	No current use
F	current use <7 drinks/week	Low level use
F	7+ drinks/week	High level use
M	no current use (never or former users)	No current use
M	current use <14 drinks/week	Low level use
M	14+ drinks/week	High level use

Response format: 0 = No current use
1 = Low level use
2 = High level use

Source variable(s):

Gender. Participant Gender

Alcohol_Use. (1=Never,2=Former,3=Current)

TotalDrinks_Per_Week. Total Weekly Alcohol Consumption

12.50 Doctor_Visit (Doctor/HCP number of visits in past 12)

This variable defines grouped categories for the number of times a participant has visited the doctor in the last 12 months. Update 12/2013: DOCTOR_VISIT was corrected to calculate values for version B for the health care use form (HCE).

```
if hcea5=0 | hceb5=0 then doctor_visit=1;
else if hcea5=1 | hceb5=1 then doctor_visit=2;
else if hcea5 in(2,3) | hceb5 in(2,3) then doctor_visit=3;
else if hcea5>=4 | hceb5>=4 then doctor_visit=4;
```

Response format: 1 = None
2 = Once
3 = 2-3 Times
4 = More than 3 times

Source variable(s):

HCEA5. During the past 12 months, how many times did you see a physician for your health care?

HCEB5. During the past 12 months, how many times did you see a physician for your health care?

12.51 Claudication_Int (Intermittent Claudication - either leg)

This is a 0/1/9 variable that defines intermittent claudication being present in either leg as follows:

```
if clea2=. then CLAUDICATION_INT=.;
else if clea2=1 and (clea7=1 | clea12=1) then CLAUDICATION_INT=1;
else if clea2=1 and (clea7=2 | clea12=2) then CLAUDICATION_INT=0;
```

else CLAUDICATION_INT=9;

Response format: 0 = Intermittent claudication not present (No)
1 = Intermittent claudication present (Yes)
9 = Don't Know

Source variable(s):

CLEA2. Do you get pain or discomfort in either leg on walking?

CLEA7. What happens to the pain if you stand still? (Right Leg)

CLEA12. What happens to the pain if you stand still? (Left Leg)

12.52 CVD_Frame (Framingham CVD composite criterion)

This is a 0/1 indicator variable that defines a composite CVD definition based on the Framingham Study criterion (see the following link for reference documentation:

<http://www.framinghamheartstudy.org/risk/gencardio.html>):

The Framingham Heart Study defines CVD as a composite of CHD (coronary death, myocardial infarction, coronary insufficiency, and angina), cerebrovascular events (including ischemic stroke, hemorrhagic stroke, and transient ischemic attack), peripheral artery disease (intermittent claudication), and heart failure].

If any of the following highlighted variables = 1, then let CVD_FRAME=1, otherwise =0. Note that this variable is meant to classify positive cases, participants with incomplete information were classified as no/unknown

CHD:

Coronary death – n/a in baseline HCHS

Myocardial infarction – MI_ECG=1 or MHEA4=1

Coronary insufficiency – considering all of the following:

Rheumatic heart disease – MHEA6=1

Atrial fibrillation- MHEA7=1

Angio/stent used to improve:

Blood flow to heart – MHEA9=1

Correct stroke – MHEA12=1

Aortic aneurysm, ballooning, or AAA – MHEA13=1

Angina – MHEA3=1

Cerebrovascular events:

including ischemic stroke - MHEA10=1

hemorrhagic stroke – MHEA10=1

transient ischemic attack - MHEA11=1

Peripheral artery disease (intermittent claudication) -

MHEA14=1 or CLAUDICATION_INT=1 (from CLEA form) or ABIGRP2_C4 = 1 (defined PAD).

Heart failure – MHEA5=1

Other heart problems - MHEA8=1

Response format: 0 = No or unknown CVD
1 = CVD risk (yes)

Source variable(s):

MHEA3. Has a doctor ever said that you have angina?

MHEA4. Has a doctor ever said that you have had a heart attack?

MHEA5. Has a doctor ever said that you had heart failure?

MHEA6. Has a doctor ever said that you had rheumatic heart disease?

MHEA7. Has a doctor ever told you that you had atrial fibrillation?

MHEA8. Has a doctor ever said that you had some other kind of heart problem?

MHEA9. Have you had a balloon angioplasty, a stent, or bypass surgery to the arteries in your heart to improve the blood flow to your heart? MHEA10. Has a doctor ever said that you had a stroke?

MHEA10. Has a doctor ever said that you had a stroke?

MHEA11. Has a doctor ever said that you had a mini-stroke of TIA (transient ischemic attack)?

MHEA12. Have you had a balloon angioplasty or surgery to the arteries of your neck to prevent or correct a stroke?

MHEA13. Has a doctor ever said that you have an aortic aneurysm, an AAA, or ballooning of your aorta?

MHEA14. Has a doctor ever said that you have peripheral arterial disease (problems with circulation, blocked arteries to the legs)?

ABIGRP2_C4. 4-level grouped Ankle Brachial Index – v2

CLAUDICATION_INT. Intermittent Claudication - either leg

MI_ECG. MI defined by ECG results

12.53 Frame_CVD_Risk_10Yr (Framingham CVD - 10 year risk (w/lab predictors))

This variable defines the Framingham CVD 10 year risk using lab predictors based on the Framingham Study criterion for participants Individuals 30 to 74 years old only (see the following link for reference documentation:

<http://www.framinghamheartstudy.org/risk/gencardio.html>):

The 10-year risk for women can be calculated as $1 - 0.95012^{\exp(\sum \beta X - 26.1931)}$ where β is the regression coefficient and X is the level for each risk factor; the risk for men is given as $1 - 0.88936^{\exp(\sum \beta X - 23.9802)}$

Diabetes definition as outlined by Framingham:

If LABA70 >= 126 OR MED_ANTIDIAB=1 OR MUEA33C=1 OR MHEA16c=1
then DIABETES_FRAME=1. Otherwise =0.

Hypertension medication use as outlined by Framingham:

If MED_ANTIHYPERT=1 OR MUEA33D=1 then HYPERT_MED_FRAME=1.
Otherwise =0.

The final definition will be based on Framingham beta coefficient and variable definitions as follows:

Men* (10-year Baseline Survival: So(10) = 0.88936)	
Variable	Beta**
Log of Age (AGE)	3.06117
Log of Total Cholesterol (LABA66)	1.12370
Log of HDL Cholesterol (LABA68)	-0.93263
Log of SBP if not treated (SBPA5 where HYPERT_MED_FRAME=0)	1.93303
Log of SBP if treated (SBPA5 where HYPERT_MED_FRAME=1)	1.99881
Smoking (Current_smoker)	0.65451
Diabetes (DIABETES_FRAME)	0.57367

Women* (10-year Baseline Survival: So(10) = 0.95012)	
Variable	Beta**
Log of Age (AGE)	2.32888
Log of Total Cholesterol (LABA66)	1.20904
Log of HDL Cholesterol (LABA68)	-0.70833
Log of SBP if not treated (SBPA5 where HYPERT_MED_FRAME=0)	2.76157
Log of SBP if treated (SBPA5 where HYPERT_MED_FRAME=1)	2.82263
Smoking (Current_smoker)	0.52873
Diabetes (DIABETES_FRAME)	0.69154

Take the log of each continuous variable as indicated above and multiply it by the indicated beta coefficient. For the categorical response variables above simply multiply the variable by the indicated beta coefficient. Then sum all of the values calculated above to define LOG_RISK, setting participants to missing if they are missing any of the above

components. Finally, calculate the desired variable, FRAME_CVD_RISK_10YR, for individuals age 30-74 only as follows:

Where $30 \leq \text{AGE} \leq 74$:

For women (GENDER='F'):

let FRAME_CVD_RISK_10YR = $1 - 0.95012^{\exp(\text{LOG_RISK} - 26.1931)}$

For men (GENDER='M'):

let FRAME_CVD_RISK_10YR = $1 - 0.88936^{\exp(\text{LOG_RISK} - 23.9802)}$

(**Note: In the Framingham definition, the beta coefficients were calculated only among those who did NOT have a previous diagnosis of CVD. This exclusion was applied to obtain the beta coefficients only and should not be, and therefore was not, used as an exclusion criterion for calculating the Framingham 10-year risk of CVD for the HCHS/SOL participants. This CVD_FRAME variable has been provided separately for potential use.)

Response format: 0 = No
1 = Yes

Source variable(s):

AGE. Age in years at the time of participant's clinic visit

LABA66. Total Cholesterol (mg/dL)

LABA68. HDL cholesterol (mg/dL)

LABA70. Fasting glucose (mg/dL)

SBPA5. Systolic blood pressure

MUEA33C. Self-report anti-diabetic medication use

MUEA33D. Self-report anti-hypertensive medication use

MHEA3. Has a doctor ever said that you have angina?

MHEA4. Has a doctor ever said that you had a heart attack?

MHEA5. Has a doctor ever said that you had heart failure?

MHEA6. Has a doctor ever said that you had rheumatic heart disease?

MHEA7. Has a doctor ever told you that you had atrial fibrillation?

MHEA8. Has a doctor ever said that you had some other kind of heart problem?

MHEA9. Have you had a balloon angioplasty, a stent, or bypass surgery to the arteries in your heart to improve the blood flow to your heart?

MHEA10. Has a doctor ever said that you had a stroke?

MHEA11. Has a doctor ever said that you had a mini-stroke or TIA?

MHEA12. Have you had a balloon angioplasty or surgery to the arteries of your neck to prevent or correct a stroke?

MHEA13. Has a doctor ever said that you have an aortic aneurysm, an AAA, or ballooning of your aorta?

MHEA14. Has a doctor ever said that you have peripheral arterial disease? MHEA16c. Are you being treated with insulin?

MED_ANTIDIAB. Anti-diabetic medication use

MED_ANTIHYPERT. Antihypertensives

CURRENT_SMOKER. Current Smoker (from CIGARETTE_USE)

12.54 HBA1C_SI (Glycosylated Hemoglobin in SI units (mmol/mol))

HBA1C_SI = (LABA72 *10.93) – 23.50, rounded to the nearest integer

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

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

Source variable(s):

LABA72. % Glycosylated Hemoglobin

12.55 CIGARETTE_PACK_YEARS_C3 (Cigarette Pack Years (3-level, cutoff =10))

Groups cigarette use into 3 categories: 0, between 0 and 10 and 10+

If CIGARETTE_PACK_YEARS =0 then set CIGARETTE_PACK_YEARS_C3=1.

If CIGARETTE_PACK_YEARS is >0 and <10 set CIGARETTE_PACK_YEARS_C3=2.

If CIGARETTE_PACK_YEARS is >=10 then set CIGARETTE_PACK_YEARS_C3=3

Response format: 1 = 0 years
2= >0 and <10 years
3= 10+ years

Source variable(s):

CIGARETTE_PACK_YEARS. Cigarette Pack Years

13. MEDICATIONS – SCANNED/TRANSCRIBED ONLY

All participants are asked to bring in all prescribed or over-the-counter medications taken in the past 4 weeks to the baseline examination for inventory. MUEA1 and MUEA2 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 MUEA5A-MUEA5D to MUEA29A-MUEA29D and MUEA39A-MUEA39D (up to 26 possible medications). For the derived indicator variables below, participants were set to missing unless the reported MUEA2=1 (took no medication) or reported at least 1 medication in the fields above.

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

1. Medi-Span’s Master Drug Data Base (MDDDB ®, 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 MDDDB ®, 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.

Hormone Replacement Therapies:

13.1 Med_Current_HRT (Any Hormone Repl. Therapy)

This is a 0/1 numeric variable that indicates use of at least one of the following HRT medications:

Med_Estrogens
Med_Estrogen_Combo
Med_Contraceptives
Med_Progestins
Med_SERM
Med_Fertility_Reg
Med_LHRH
Med_LHRH_AG
Med_LHRH_ANT

13.2 Med_Estrogens (Estrogens)

This is a 0/1 numeric variable.

Medication Codes:

The "24xxxx" group of medications: *ESTROGENS*

13.3 Med_Estrogen_Combo (Estrogens – Combinations Only)

This is a 0/1 numeric variable.

Medication Codes:

The "2499xx" class of medications: *Estrogen Combinations**

13.4 Med_Contraceptives (Contraceptives)

This is a 0/1 numeric variable.

Medication Codes:

The "25xxxx" group of medications: *CONTRACEPTIVES*

13.5 Med_Progestins (Progestins)

This is a 0/1 numeric variable.

Medication Codes:

The "26xxxx" group of medications: *PROGESTINS*

13.6 Med_SERM (Selective Estrogen Receptor Modulators)

This is a 0/1 numeric variable.

Medication Codes:

The “300530” subclass of medications: *Selective Estrogen Receptor Modulators (SERMs)***

13.7 Med_Fertility_Reg (Fertility Regulators)

This is a 0/1 numeric variable.

Medication Codes:

The “3006xx” class of medications:

13.8 Med_LHRH (LHRH/GnRH Only)

This is a 0/1 numeric variable.

Medication Codes:

The “300700” class of medications: *Luteinizing Hormone Releasing-Hormones (LHRH/GnRH)**

13.9 Med_LHRH_AG (LHRH/GnRH Agonist Analog)

This is a 0/1 numeric variable.

Medication Codes:

The “300800” subclass of medications: *LHRH/GnRH Agonist Analog Pituitary Suppressants**

13.10 Med_LHRH_ANT (LHRH/GnRH Antagonists)

This is a 0/1 numeric variable.

Medication Codes:

The “300900” subclass of medications: *GnRH/LHRH Antagonists**

Other Medications:

13.11 Med_Antiangular (Antianginals)

This is a 0/1 numeric variable.

Medication Codes:

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

13.12 Med_Antiarrhythmics (Antiarrhythmics)

This is a 0/1 numeric variable.

Medication Codes:

The “35xxxx” group of medications: *ANTIARRHYTHMICS*

13.13 Med_Antiasthmatics (Antiasthmatic or bronchodilator agents)

This is a 0/1 numeric variable.

Medication Codes:

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

13.14 Med_Anticoag (Anticoagulants)

This is a 0/1 numeric variable.

Medication Codes:

Any “83xxxx” group of medications: *ANTICOAGULANTS*

13.15 Med_Anticoag_Coumarin (Coumarin anticoagulants)

This is a 0/1 numeric variable.

Medication Codes:

Any “8320xx” class of medications: *Coumarin Anticoagulants**

13.16 Med_Anticoag_Heparin (Heparin anticoagulants)

This is a 0/1 numeric variable.

Medication Codes:

Any “8310xx” class of medications: *Heparins And Heparinoid-Like Agents**

13.17 Med_Antidiab (Antidiabetics)

This is a 0/1 numeric variable.

Medication Codes:

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

13.18 Med_Antihypert (Antihypertensives)

This is a 0/1 numeric variable.

Medication Codes:

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

13.19 Med_Antihypert_ACEI (ACE Inhibitors)

This is a 0/1 numeric variable.

Medication Codes:

Any "3610xx" class of medications: *ACE Inhibitors**

Any "369915" subclass of medications: *ACE Inhibitor & Calcium Channel Blocker Combinations***

Any "369918" subclass of medications: *ACE Inhibitors & Thiazide/Thiazide-Like***

13.20 Med_Antihypert_AT2RAs (Angiotension II Receptor Antagonists)

This is a 0/1 numeric variable.

Medication Codes:

Any "361500" subclass of medications: *Angiotensin II Receptor Antagonists**

Any "369940" subclass of medications: *Angiotensin II Receptor Antagonists & Thiazides***

13.21 Med_Antiplatelet (Antiplatelets)

This is a 0/1 numeric variable.

Medication Codes:

Any "8515xx" class of medications: *Platelet Aggregation Inhibitors**

13.22 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**

13.23 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***

12/2019 update: Any "369988" subclass of medications: *Beta Blocker-Nutritional supplement combinations***

13.24 Med_BB_Ophthalmic (Ophthalmic Beta Blockers)

This is a 0/1 numeric variable.

Medication Codes:

Any "862500" subclass of medications: *Beta-blockers - Ophthalmic**

Any "862599" subclass of medications: *Beta-blockers - Ophthalmic Combinations***

13.25 Med_CACB (Calcium Channel Blockers)

This is a 0/1 numeric variable.

Medication Codes:

Any "34xxxx" group of medications:

Any "369915" subclass of medications: *CALCIUM CHANNEL BLOCKERS*

Any "369925" subclass of medications: *Beta Blocker & Calcium Channel Blocker Combinations***

Any "409925" subclass of medications: *Calcium Channel Blocker & HMG CoA Reductase Inhibit Comb***

12/2019 update: Any "369930" subclass of medications: *Angiotensin II Receptor Antagonist & Calcium Channel Blocker Combinations***

13.26 Med_Cardiacycosides (Cardiac Glycosides)

This is a 0/1 numeric variable.

Any "312000" subclass of medications: *Cardiac Glycosides**

13.27 Med_Chemo (Chemotherapy)

This is a 0/1 numeric variable.

Medication Codes:

Any "21xxxx" group of medications: *ANTINEOPLASTICS*

13.28 Med_Clopid (Clopidogrel (Plavix))

This is a 0/1 numeric variable.

Medication Codes:

The "851580201" subclass of medications: *Thienopyridine Derivatives***

13.29 Med_COX2 (COX-2 inhibitors)

This is a 0/1 numeric variable.

Medication Codes:

Any "661005" subclass of medications: *Cyclooxygenase 2 (COX-2) Inhibitors***

13.30 Med_Diuretic (Diuretics)

This is a 0/1 numeric variable.

Medication Codes:

Any "369918" subclass of medications: *ACE Inhibitors & Thiazide/Thiazide-Like***

Any "369920" subclass of medications: *Beta Blocker & Diuretic Combinations***

Any "369940" subclass of medications: *Angiotensin II Receptor Antagonists & Thiazides***

Any "369950" subclass of medications: *Adrenolytics-Central & Thiazide Combinations***

Any "369955" subclass of medications: *Adrenolytics-Peripheral & Thiazides***

Any "369970" subclass of medications: *Antihypertensives-MAOIs & Thiazides***

Any "369980" subclass of medications: *Antihypertensives-MISC & Thiazides***

Any "369990" subclass of medications: *Vasodilators & Thiazides***

Any "37xxxx" group of medications: *DIURETICS*

13.31 Med_Diuretic_Thiazide (Thiazide Diuretics)

This is a 0/1 numeric variable.

Medication Codes:

Any "369918" subclass of medications: *ACE Inhibitors & Thiazide/Thiazide-Like***

Any "369920" subclass of medications: *Beta Blocker & Diuretic Combinations***

Any "369940" subclass of medications: *Angiotensin II Receptor Antagonists & Thiazides***

Any "369950" subclass of medications: *Adrenolytics-Central & Thiazide Combinations***

Any "369955" subclass of medications: *Adrenolytics-Peripheral & Thiazides***

Any "369970" subclass of medications: *Antihypertensives-MAOIs & Thiazides***

Any "369980" subclass of medications: *Antihypertensives-MISC & Thiazides***

Any "369990" subclass of medications: *Vasodilators & Thiazides***

Any "376000" subclass of medications: *Thiazides and Thiazide-Like Diuretics**

13.32 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***

13.33 Med_Insulin (Insulins)

This is a 0/1 numeric variable.

Medication Codes:

Any "2710xx" class of medications: *Insulin**

13.34 Med_LLD (Lipid lowering drugs/Antihyperlipidemics (LLD))

This is a 0/1 numeric variable.

Medication Codes:

Any "39xxxx" group of medications: *ANTIHYPERLIPIDEMICS* plus the class of "409925" (*Calcium Channel Blocker & HMG CoA Reductase Inhibit Comb***).

13.35 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**

13.36 Med_NSAID (NSAIDs)

This is a 0/1 numeric variable.

Medication Codes:

Any "6610xx" class of medications: *Nonsteroidal Anti-inflammatory Agents (NSAIDs)**

13.37 Med_OI_Steroid (Oral/inhalable Glucocorticosteroids)

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

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***

13.38 Med_Statin (Statins)

This is a 0/1 numeric variable.

Medication Codes:

Any "3940xx" class of medications HMG CoA Reductase Inhibitors

12/2019 update: Any "399940" subclass of medications: *Intestinal Cholesterol Absorption Inhibitor-HMG CoA Reductase Inhibitor Combinations***

12/2019 update: Any "409925" subclass of medications: *Calcium Channel Blocker & HMG CoA Reductase Inhibitor Combinations***

13.39 Med_ANTIAXI (Antianxiety agents)

This is a 0/1 numeric variable.

Medication Codes:

Any "57xxxx" class of medications: "ANTIAXIETY AGENTS"

13.40 Med_ANTIDEPRESS (Antidepressants)

This is a 0/1 numeric variable.

Medication Codes:

Any "58xxxx" class of medications: "ANTIDEPRESSANTS"

13.41 Med_ANTIDEPRE_SSRI (Antidepressants - selective serotonin re-uptake inhibitors)

This is a 0/1 numeric variable.

Medication Codes:

Any "5816xx" class of medications: "SELECTIVE SEROTONIN REUPTAKE INHIBITORS"

13.42 Med_ANTIDEPRE_TRICYC (Antidepressants - tricyclic antidepressants)

This is a 0/1 numeric variable.

Medication Codes:

Any "5820xx" class of medications: "TRICYCLIC AGENTS"

13.43 Med_ANTIPSYCHO (Antipsychotics)

This is a 0/1 numeric variable.

Medication Codes:

Any "59xxxx" class of medications: "ANTIPSYCHOTICS"

13.44 Med_CORTICOSTEROIDS_INHALED (Inhaled Corticosteroid medication)

This is a 0/1 numeric variable.

Medication Codes:

The "221000" subclass of medications: "GLUCOCORTICOSTEROIDS"

The "221099" subclass of medications: "STEROID COMBINATIONS"

13.45 Med_PDE4 (Phosphodiesterase IV Inhibitor Medication)

This is a 0/1 numeric variable.

Medication Codes:

Any "311000" subclass of medications: "PHOSPHODIESTERASE INHIBITORS"

13.46 Med_BRONCHODILATOR (Bronchodilator and Antiasthmatic Agents medication)

This is a 0/1 numeric variable.

Medication Codes:

The "441000" subclass of medications: "BRONCHODILATORS - ANTICHOLINERGICS"

The "442010" subclass of medications: "BETA ADRENERGICS"

The "442020" subclass of medications: "MIXED ADRENERGICS"

The "443000" subclass of medications: "XANTHINES"
 The "444000" subclass of medications: "STEROID INHALANTS"
 The "445000" subclass of medications: "LEUKOTRIENE MODULATORS"
 The "445040" subclass of medications: "5-LIPOXYGENASE INHIBITORS"
 The "445050" subclass of medications: "LEUKOTRIENE RECEPTOR ANTAGONISTS"
 The "446000" subclass of medications: "ANTIASTHMATIC - MONOCLONAL ANTIBODIES"
 The "446030" subclass of medications: "ANTI-IGE MONOCLONAL ANTIBODIES"
 The "449900" subclass of medications: "ASTHMA AND BRONCHODILATOR AGENT COMBINATIONS"
 The "449905" subclass of medications: "XANTHINE MIXTURES"
 The "449910" subclass of medications: "XANTHINE - EXPECTORANTS"
 The "449920" subclass of medications: "XANTHINE - SYMPATHOMIMETICS"
 The "449922" subclass of medications: "XANTHINE – SYMPATHOMIMETIC - EXPECTORANT"
 The "449925" subclass of medications: "XANTHINE - BARBITURATES"
 The "449927" subclass of medications: "SYMPATHOMIMETIC - BARBITURATE"
 The "449928" subclass of medications: "XANTHINE-ANTIHISTAMINE-EXPECTORANT"
 The "449930" subclass of medications: "XANTHINE – SYMPATHOMIMETIC - BARBITURATE"
 The "449931" subclass of medications: "Xanthine-Sympathomimetic-Barbiturate-Antihistamine"
 The "449932" subclass of medications: "XANTHINE – SYMPATHO – BARBIT - EXPECTOR"
 The "449950" subclass of medications: "XANTHINE - ANTITUSSIVE"
 The "449952" subclass of medications: "XANTHINE-ANTITUSSIVE-ANTIHISTAMINE"
 The "449955" subclass of medications: "XANTHINE-ANTITUSSIVE-EXPECTORANT"
 The "449985" subclass of medications: "XANTHINE-ANTIHISTAMINES"
 The "449988" subclass of medications: "Xanthine-Steroids"
 The "449990" subclass of medications: "Xanthine-Sympathomimetic-Antihistamine-~/or Expectorant"
 The "449993" subclass of medications: "Xanthine-Sympathomimetic-Antitussive-Expectorant"
 The "449995" subclass of medications: "Xanthine-Sympath-Antihistamine-Antitussive-Expectorant"

13.47 Med_PULMONARY (Pulmonary medication)

This is a 0/1 numeric variable.

Medication Codes:

The "221000" subclass of medications: "GLUCOCORTICOSTEROIDS"

The "221099" subclass of medications: "STEROID COMBINATIONS"

The "311000" subclass of medications: "PHOSPHODIESTERASE INHIBITORS"

The "441000" subclass of medications: "BRONCHODILATORS - ANTICHOLINERGICS"

The "442010" subclass of medications: "BETA ADRENERGICS"

The "442020" subclass of medications: "MIXED ADRENERGICS"

The "443000" subclass of medications: "XANTHINES"

The "444000" subclass of medications: "STEROID INHALANTS"

The "445000" subclass of medications: "LEUKOTRIENE MODULATORS"

The "445040" subclass of medications: "5-LIPOXYGENASE INHIBITORS"

The "445050" subclass of medications: "LEUKOTRIENE RECEPTOR ANTAGONISTS"

The "446000" subclass of medications: "ANTIASTHMATIC - MONOCLONAL ANTIBODIES"

The "446030" subclass of medications: "ANTI-IGE MONOCLONAL ANTIBODIES"
The "449900" subclass of medications: "ASTHMA AND BRONCHODILATOR AGENT COMBINATIONS"
The "449905" subclass of medications: "XANTHINE MIXTURES"
The "449910" subclass of medications: "XANTHINE - EXPECTORANTS"
The "449920" subclass of medications: "XANTHINE - SYMPATHOMIMETICS"
The "449922" subclass of medications: "XANTHINE – SYMPATHOMIMETIC - EXPECTORANT"
The "449925" subclass of medications: "XANTHINE - BARBITURATES"
The "449927" subclass of medications: "SYMPATHOMIMETIC - BARBITURATE"
The "449928" subclass of medications: "XANTHINE-ANTIHISTAMINE-EXPECTORANT"
The "449930" subclass of medications: "XANTHINE – SYMPATHOMIMETIC - BARBITURATE"
The "449931" subclass of medications: "Xanthine-Sympathomimetic-Barbiturate-Antihistamine"
The "449932" subclass of medications: "XANTHINE – SYMPATHO – BARBIT - EXPECTOR"
The "449950" subclass of medications: "XANTHINE - ANTITUSSIVE"
The "449952" subclass of medications: "XANTHINE-ANTITUSSIVE-ANTIHISTAMINE"
The "449955" subclass of medications: "XANTHINE-ANTITUSSIVE-EXPECTORANT"
The "449985" subclass of medications: "XANTHINE-ANTIHISTAMINES"
The "449988" subclass of medications: "Xanthine-Steroids"
The "449990" subclass of medications: "Xanthine-Sympathomimetic-Antihistamine-&/or Expectorant"
The "449993" subclass of medications: "Xanthine-Sympathomimetic-Antitussive-Expectorant"
The "449995" subclass of medications: "Xanthine-Sympath-Antihistamine-Antitussive-Expectorant"

14. MEDICATIONS – OTHER

14.1 Asthma_Meds (Asthma Med Use – self/scanned (curr asthma by MD diag))

This is a 0/1 numeric variable. This variable combines scanned/transcribed anti-asthmatic medication use with self-report asthma medication use. This variable is defined as 0 for those that do not currently have asthma based on an MD diagnosis.

If ASTHMA_CURR_MD=0 then ASTHMA_MEDS=0.
Else if ASTHMA_CURR_MD=1 and (RSEA37=1 or MED_ANTIASTHMATICS=1)
then ASTHMA_MEDS=1.
Else if ASTHMA_CURR_MD=1 and (RSEA37=0 or MED_ANTIASTHMATICS=0)
then ASTHMA_MEDS=0.
Otherwise, ASTHMA_MEDS=missing.

Response format: 0 = No current asthma or no anti-asthmatic medication use
1 = Anti-asthmatic medication use

Source Variable(s):

RSEA37. Ever take asthma medication in last 12 months.
ASTHMA_CURR_MD. Currently have asthma with ever MD diagnosis
MED_ANTIASTHMATICS. Anti-asthmatic Medication Use

14.2 COPD_Meds (COPD Med Use – self/scanned (ever COPD by MD diag))

This is a 0/1 numeric variable. This variable combines scanned/transcribed anti-asthmatic medication use with self-report COPD medication use. This variable is defined as 0 for those that have never had COPD/emphysema or chronic bronchitis based on an MD diagnosis.

If COPD_EVER=1 and (RSEA46=1 or MED_ANTIASTHMATICS=1) then
COPD_MEDS=1.
Else if COPD_EVER = non-missing then COPD_MEDS=0.
Otherwise, COPD_MEDS=missing.

Response format: 0 = Never had COPD/emphysema or no indication of COPD/anti-asthmatic medication use
1 = COPD/anti-asthmatic medication use

Source Variable(s):

RSEA46. Ever take COPD/Emph medication in last 12 months.
COPD_EVER. Ever had COPD/Emph or CB
MED_ANTIASTHMATICS. Anti-asthmatic Medication Use

15. NEUROCOGNITIVE

15.1 ECG_MAJ (Major ECG Abnormalities)

Defined only for participants that have ECG data.

First, derived corresponding Binary indicator derived variables: ECG_266 – ECG_280:

For the 15 source variables (from ECGA266 to ECGA280):

If the variable has missing value (no text), then set a Binary indicator variable as 0;

If the variable has descriptive text as the value, then set a Binary indicator variable as 1.

Second, derive the count variables for the major ECG abnormalities:

ECG_Maj_n (count of Major ECG Abnormalities) = the sum of Binary indicator variables ECG_266 – ECG_280

Then define the Major ECG abnormalities variables:

If ECG_Maj_N > 0 then set Major ECG abnormalities ECG_Maj as 1, otherwise if ECG_Maj_N = 0 then set ECG_Maj as 0;

Response format: 0 = No
1 = Yes

Source Variable(s):

ECGA266. one of the following

ECGA267. Major Q Wave abnormalities [old Myocardial Infarction (MI)]

ECGA268. Minor Q, QS waves with ST, T abnormalities (possible old MI)

ECGA269. Major Isolated_ST_T_abnormalities

ECGA270. Left Ventricular Hypertrophy

ECGA271. Atrial Fibrillation or Flutter

ECGA272. one of the following

ECGA273. Complete (third degree) A-V block

ECGA274. one of the following

ECGA275. Wolff-Parkinson-White pattern

ECGA276. Major QT Prolongation, QT_i ≥ 116%

ECGA277. Pacemaker

ECGA278. one of the following

ECGA279. one of the following

ECGA280. Supraventricular tachycardia intermittent

15.2 ECG_MIN (Minor ECG Abnormalities)

Defined only for participants that have ECG data.

First, derived corresponding Binary indicator derived variables: ECG_281 – ECG_300:

For the 20 source variables (from ECGA281 to ECGA300):

If the variable has missing value (no text), then set a Binary indicator variable as 0;

If the variable has descriptive text as the value, then set a Binary indicator variable as 1.

Second, derive the count variables for the minor ECG abnormalities:

ECG_Min_n (count of Minor ECG Abnormalities) = the sum of Binary indicator variables
ECG_281 – ECG_300

Then define the Minor ECG abnormalities variable:

If ECG_Min_N > 0 then set Minor ECG abnormalities ECG_Min as 1, otherwise if
ECG_Min_N =0 then set ECG_Min as 0;

Response format: 0 = No
1 = Yes

Source Variable(s):

ECGA281. MINOR ISOLATED Q, QS waves
ECGA282. MINOR ISOLATED ST, T Abnormalities
ECGA283. HIGH R waves
ECGA284. ST Segment Elevation, Anterolateral Site
ECGA285. ST Segment Elevation, Posterior Site
ECGA286. ST Segment Elevation, Anterior Site
ECGA287. Incomplete RBBB
ECGA288. Incomplete LBBB
ECGA289. Minor QT Prolongation, QT_i >=112%
ECGA290. Short PR
ECGA291. Left axis deviation
ECGA292. Right axis deviation
ECGA293. Frequent Ventricular Premature Beats
ECGA294. P-R Interval > .22 sec
ECGA295. one of the following
ECGA296. Supraventricular rhythm
ECGA297. Sinus Tachycardia
ECGA298. Sinus Bradycardia
ECGA299. Low QRS amplitude
ECGA300. P-wave amplitude high

15.3 No_Disability_6Item (No Disability on NEEA5-7 (1=Yes,0=No))

This is a 0/1 numeric variable..

1, if NEEA1 in (1,2,3);
0, if NEEA1 not in (1,2,3), excluding permanently missing forms and forms with all fields marked "unresolvable";
Missing otherwise.

5/2014 update: revised to set participants missing NEE forms to missing

Response format: 0 = Disability
1 = No disability

Source Variable(s):

NEEA1. Number of presentations necessary for the participant to repeat the words.

15.4 SEVLT_3Trials (SEVTL, Sum of 3 Trials (NEEA8=10))

This is a continuous variable for the sum of the scores of the first 3 trials of the SEVLT test. Only defined for participants who took a neurocognitive assessment (NEE form).

$SEVLT_3trials = NEEA8 + NEEA9 + NEEA10$
SEVLT_3trials is missing if any of NEEA8, NEEA9, NEEA10 is missing.

Source variable(s):

NEEA8. Words Recalled from Part B (Trial 1).
NEEA9. Words Recalled from Part B (Trial 2).
NEEA10. Words Recalled from Part B (Trial 3).

15.5 SEVLT_Recall (SEVLT, Delay Recall (NEEA11 with re-coding))

This is a continuous variable for the scores of the 5th delayed recall trial of the SEVLT test. Definition per Dr. Hector Gonzalez. Only defined for participants who took a neurocognitive assessment (NEE form).

$SEVLT_recall = 0$, if NEEA11 is missing and SEVLT_3trials is not missing,
Otherwise $SEVLT_recall = NEEA11$.

Source variable(s):

NEEA11. Words Recalled from Part B (Trial 5).
SEVLT_3Trials. SEVTL, Sum of 3 Trials (NEEA8=10)

15.6 Total_6Item (Total Score on 6 Item Screener (with re-coding))

This is a continuous variable for the total score of the 6-item screener, which uses recoding of the original responses per definitions by Dr. Hector Gonzalez. It can take integer values from 0 to 6. Only defined for participants who took a neurocognitive assessment (NEE form).

The computation of recoded variables is according to the table below:

Variable	Definition
No_Disability_6item	1, if NEEA1 in (1,2,3); 0, if NEEA1 not in (1,2,3), excluding permanently missing forms and forms with all fields marked "unresolvable"; Missing otherwise.
NEEA2_r1	=2, if NEEA2 in (2,3,4); = NEEA2 otherwise.
NEEA3_r1	=2, if NEEA3 in (2,3,4); = NEEA3 otherwise.
NEEA4_r1	=2, if NEEA4 in (2,3,4); = NEEA4 otherwise.
NEEA5_r1	=2, if NEEA5 in (2,3,4) and No_Disability_6item=1; = NEEA5 otherwise.
NEEA6_r1	=2, if NEEA6 in (2,3,4) and No_Disability_6item=1; = NEEA6 otherwise.
NEEA7_r1	=2, if NEEA7 in (2,3,4) and No_Disability_6item=1; = NEEA7 otherwise.
NEEA2_recode	= 0, if NEEA2_r1=2; =1, if NEEA2_r1=1; = missing, if NEEA2_r1 in (3,4) or missing.
NEEA3_recode	= 0, if NEEA3_r1=2; =1, if NEEA3_r1=1; = missing, if NEEA3_r1 in (3,4) or missing.
NEEA4_recode	= 0, if NEEA4_r1=2; =1, if NEEA4_r1=1; = missing, if NEEA4_r1 in (3,4) or missing.
NEEA5_recode	= 0, if NEEA5_r1=2; =1, if NEEA5_r1=1; = missing, if NEEA5_r1 in (3,4) or missing.
NEEA6_recode	= 0, if NEEA6_r1=2; =1, if NEEA6_r1=1; = missing, if NEEA6_r1 in (3,4) or missing.
NEEA7_recode	= 0, if NEEA7_r1=2; =1, if NEEA7_r1=1; = missing, if NEEA7_r1 in (3,4) or missing.

Total_6item = NEEA2_recode + NEEA3_recode + NEEA4_recode + NEEA5_recode +
NEEA6_recode + NEEA7_recode

Total_6item should be missing if any of the variables in the sum is missing.

Source variable(s):

NEEA2. "What year is this?"

NEEA3. "What month is this?"

NEEA4. "What is the day of the week?"

NEEA5. "Now, what were those three words I asked you to remember?" – Blue

NEEA6. "Now, what were those three words I asked you to remember?" – Pear

NEEA7. "Now, what were those three words I asked you to remember?" – Sofa

No_Disability_6Item. No Disability on NEEA5-7 (1=Yes,0=No)

15.7 WORDFREQ (Word Frequency (A+F) (NEEA12-13))

This is a continuous variable for the sum of the score of the A and F parts of the Word Fluency test. Only defined for participants who took a neurocognitive assessment (NEE form).

=NEEA12 + NEEA13

- should be missing if any of them is missing.

Source Variable(s):

NEEA12. Words Produced on Part C: Letter F

NEEA13. Words Produced on Part C: Letter A

15.8 DIGITSYMBOL (Digit Symbol Substitution Score(NEEA14))

This is a continuous variable for the score of the digit symbol substitution test. Only defined for participants who took a neurocognitive assessment (NEE form).

DigitSymbol = NEEA14

Source variable(s):

NEEA14. Total Correct Symbols on PART D: DIGIT SYMBOL SUBSTITUTION.

16. DENTAL

16.1 Periodontal_Disease (Periodontal Disease - NIDCR)

This is a 0/1 numeric variable. The NIDCR cites the prevalence of periodontal disease based on dental examination probing measurements. A person must have at least one site with 3mm or more of attachment loss and 4mm or more of pocket depth in order to be positive for periodontal disease.

Periodontal_Disease = there is at least one tooth with Attachment Loss (AL) 3mm or greater AND Pocket Depth (PD) 4mm or greater.

Response format: 0 = No periodontal disease (No)
1 = Periodontal disease detected (Yes)

```
** pocket depth (PD) measures;  
array pd(6,32) perapmb1-perapmb32  
    perapbu1-perapbu32  
    perapdb1-perapdb32  
    perapdl1-perapdl32  
    perapln1-perapln32  
    perapml1-perapml32;  
array pdx(32) pdx1-pdx32;
```

```
** attachment loss (AL) measures = PD - CEJ (must be >= 0);  
array al(6,32) almb1-almb32  
    albu1-albu32  
    aldb1-aldb32  
    aldl1-aldl32  
    alln1-alln32  
    alml1-alml32;  
array alx(32) alx1-alx32;
```

```
array chx(32) chx1-chx32;
```

```
do k=1 to 32;  
  do l=1 to 6;  
    if pd(l,k) >=0 then do;  
      if pd(l,k) >=4 then pdx(k)=1;  
    end;  
  
    if al(l,k) >=-100 then do;  
      if al(l,k) >=3 then alx(k)=1;  
    end;  
  end;  
end;
```

```
if pdx(k)=1 & alx(k)=1 then chx(k)=1;
```

```
else chx(k)=0;
end;
```

```
if sum(of chx1-chx32)>=1 then PERIODONTAL_DISEASE=1;
else PERIODONTAL_DISEASE=0;
```

Source variable(s):

Pocket Depth and Attachment loss variable arrays from Dental Derive File (see above code).

16.2 PERIO_CDC_AAP (Periodontitis Severity using CDC/AAP Grades)

We want to extend the definition of periodontal disease to include a derived variable that uses the most recent Centers for Disease Control (CDC) and American Academy of Periodontology (AAP) severity grading guidelines. These definitions are different than the existing one for PERIODONTAL_DISEASE in the current PART_DERV file. The following specifications were provided by Jim Beck at the UNC Dental School. Clinical measures to assess periodontal health include probing pocket depth (PD) and gingival recession; where PD refers to the depth of the pocket measured from the gingival margin (GM) to the depth of the gingival sulcus, and gingival recession is the distance from the GM to the cemento-enamel junction (CEJ). Clinical attachment level (CAL) is calculated from the sum of PD and GM-CEJ measurements and forms the basis for the currently accepted periodontal disease classification system. For attachment loss, (AL) measures = PD - CEJ (must be ≥ 0).

The HCHS/SOL used a comprehensive full mouth examination protocol examining six sites per tooth in up to 28 teeth yielding up to 168 sites per person. For the purpose of this study, four outcome measures will be reported based on the Centers for Disease Control and Prevention and the American Academy of Periodontology (CDC/AAP) case definition of periodontitis at four severity thresholds: (1) mild; (2) moderate; (3) severe; (4) total. These classifications (described below) are designed for population-based surveillance of periodontal disease and are based on measures of clinical periodontal clinical attachment level (CAL) and probing pocket depth (PD) thus defined:

Mild: [2 or more interproximal sites with 3 mm or more CAL AND 2 or more sites with 4 mm or more PD] (not on same tooth)
OR
[At least 1 site with 5mm or more PD]

Moderate: [2 or more interproximal sites with CAL 4 mm or more (not on same tooth)]
OR
[2 or more interproximal sites with PD 5 mm or more (not on same tooth)]

Severe: [2 or more interproximal sites with CAL 6 mm or more (not on same tooth)]

AND

[At least 1 tooth with an interproximal site with PD of 5 mm or more]

Response format: . = Missing if all values for pocket depth and attachment loss are missing
0 = Normal
1 = Mild periodontitis
2 = Moderate periodontitis
3 = Severe periodontitis

16.3 PERIO_Extent (Percentage Interproximal Sites Positive for Periodontitis)

This is an additional method of describing the extent of disease in the mouth is the Extent of disease, which is the percent of sites within an individual that have a CAL score ≥ 3 mm or a PD score $\geq 4, 5, \text{ or } 6$ mm. Each tooth has 4 potential measurement sites for evaluation so the individual denominator for extent of disease would be (4 x number of teeth present). The measures taken on the buccal and the lingual sites are not used in this scoring method so the array variables below have those variable ranges deleted for the buccal and lingual measures on all 32 tooth positions.

PERIO_EXTENT = (Sum positive sites per person / Number of possible sites) * 100
Note the number of possible sites = 4 * (total tooth count)

- a. Attachment loss and the CEJ and pocket depth measurements as follows:

Pocket depths:

perapmbu1 to perapmb32
~~perapbu1-perapbu32~~
perapdb1-perapdb32
perapdl1-perapdl32
~~perapln1-perapln32~~
perapml1-perapml32;

Cemento-enamel junction (CEJ) measurements:

peracmb1 to peracmb32
~~peracbu1-peracbu32~~
peracdb1-peracdb32
peracd11-peracd132
~~peracln1-peracln32~~
peracml1-peracml32;

Attachment loss:

almb1 to almb32
~~albu1-albu32~~
aldb1-aldb32
aldl1-aldl32
~~alln1-alln32~~
alml1-alml32

Note that the “bu1-bu32”, and the “ln1-ln32” rows in the arrays above are not used in the CDC definition of periodontal disease.

The strike-through character above means ignore these sites when evaluation is done for the CDC severity scoring method.

17. SLEEP

17.1 ESS (Epworth Sleepiness Scale)

This is an integer value with range from 0-24. The Epworth Sleepiness Scale assesses the daytime sleepiness of patients. Dr Murray W Johns owns the copyright in the ESS (Copyright © M.W.Johns 1990-1997). The official website is <http://epworthsleepinessscale.com/>

Algorithm: The ESS is computed by adding ALL 8 ITEMS (SLEA12A TO SLEA12H). If any item is missing then ESS is missing.

Warning: The levels of each item need to be recoded as follows: No chance (from 1 to 0), slight chance (from 2 to 1), moderate chance (from 3 to 2) and high chance (from 4 to 3).

Source variable(s):

What is the chance that you would doze off or fall asleep (not just "feel tired") in each of the following situations?

SLEA12a. Sitting and reading

SLEA12b. Watching TV

SLEA12c. Sitting inactive in a public place (such as a theater or a meeting)

SLEA12d. Riding as a passenger in a car for an hour without a break

SLEA12e. Lying down to rest in the afternoon when circumstances permit

SLEA12f. Sitting and talking to someone

SLEA12g. Sitting quietly after a lunch without alcohol

SLEA12h. In a car, while stopped for a few minutes in traffic

17.2 ESS_GT10 (Excessive sleepiness (ESS > 10))

This is a 0/1 numeric variable. If ESS is non-missing, then:

If $ESS > 10$ then $ESS_GT10 = 1$.

Otherwise if $0 \leq ESS \leq 10$ then $ESS_GT10 = 0$.

REFERENCE: Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*. 1991 Dec;14(6):540-5. PMID: [1798888](#)

Source variable(s):

ESS. Epworth Sleepiness Scale

17.3 AHI_GE15 (Sleep Apnea defined as Apnea/Hypopnea Index (3% desat) ≥ 15)

This is a 0/1 numeric variable. If SLPA54 is non-missing, then:

If $SLPA54 \geq 15$ then $AHI_GE15 = 1$.

Otherwise if $0 \leq SLPA54 < 15$ then $AHI_GE15 = 0$

Source variable(s):

SLPA54. Apnea/Hypopnea Index (3% desat)

17.4 SBD5 (Sleep Apnea Syndrome (Apnea/Hypopnea Index (3% desat) >= 5 & ESS >=10))

This is a 0/1 numeric variable. If SLPA54 and ESS are non-missing, then:
If SLPA54>=5 & ESS>=10 then SBD5=1
Otherwise SBD5=0.

Source variable(s):

SLPA54. Apnea/Hypopnea Index (3% desat)
ESS. Epworth Sleepiness Scale

17.5 SLPDUR_WKDAY (Weekday sleep duration (hours))

This is a numeric variable that calculates the self-reported sleep duration during weekdays. Sleep data was cleaned for creation of sleep duration variables. Sleep durations less than 3 hours or greater than 14 hours were set to missing.

```
IF SLEA2A_2401 <= .z OR SLEA1A_2401 <= .z then SLPDUR_WKDAY = .;  
ELSE IF SLEA2A_2401 > .z AND SLEA1A_2401 > .z then do;  
  IF SLEA2A_2401 >= SLEA1A_2401  
    then SLPDUR_WKDAY = (SLEA2A_2401 - SLEA1A_2401)/(60*60);  
  ELSE IF SLEA2A_2401 < SLEA1A_2401  
    then SLPDUR_WKDAY = (SLEA2A_2401+86400 -  
  SLEA1A_2401)/(60*60);  
end;
```

Source variable(s):

SLEA1A_2401. What time do you usually go to bed? (on weekdays)
SLEA2A_2401. What time do you usually wake up? (on weekdays)

17.6 SLPDUR_WKEND (Weekend sleep duration (hours))

This is a numeric variable that calculates the self-reported sleep duration during weekend days. Sleep data was cleaned for creation of sleep duration variables. Sleep durations less than 3 hours or greater than 14 hours were set to missing.

```
IF SLEA2C_2401 <= .z OR SLEA1C_2401 <= .z then SLPDUR_WKEND = .;  
ELSE IF SLEA2C_2401 > .z AND SLEA1C_2401 > .z then do;  
  IF SLEA2C_2401 >= SLEA1C_2401  
    then SLPDUR_WKEND = (SLEA2C_2401 - SLEA1C_2401)/(60*60);  
  ELSE IF SLEA2C_2401 < SLEA1C_2401  
    then SLPDUR_WKEND = (SLEA2C_2401+86400 -  
  SLEA1C_2401)/(60*60);  
end;
```

Source variable(s):

SLEA1C_2401. What time do you usually go to bed? (on weekends)

SLEA2C_2401. What time do you usually wake up? (on weekends)

17.7 SLPDUR (Average sleep duration (hours))

This is a numeric variable that calculates the weighted average daily sleep duration during an entire week (five weekdays and two weekend days), using self-reported sleep data. Sleep data was cleaned for creation of sleep duration variables. Sleep durations less than 3 hours or greater than 14 hours were set to missing.

$$\text{SLPDUR} = (\text{SLPDUR_WKEND} * 2 + \text{SLPDUR_WKDAY} * 5) / 7;$$

Source variable(s):

SLPDUR_WKDAY. Weekday sleep duration (hours)

SLPDUR_WKEND. Weekend sleep duration (hours)

17.8 MID_SLEEP_TIME_WKDAY (Midpoint sleep time in clock hours; weekdays)

This is a numeric variable that calculates the midpoint of self-reported sleep time during weekdays. It is in clock hours in decimal point (e.g. 3.5 is 3:30am and 15.75 is 5:45pm). Sleep data was cleaned for creation of midpoint sleep time variables.

```
IF SLEA2A_2401 <= .z OR SLEA1A_2401 <= .z then MID_SLEEP_TIME_WKDAY = .;
  ELSE IF SLEA2A_2401 > .z AND SLEA1A_2401 > .z then do;
IF SLEA2A_2401 >= SLEA1A_2401
  then MID_SLEEP_TIME_WKDAY = ((SLEA2A_2401 + SLEA1A_2401)/2) / (60*60);
ELSE IF SLEA2A_2401 < SLEA1A_2401
  then MID_SLEEP_TIME_WKDAY = ((SLEA2A_2401+86400 + SLEA1A_2401)/2)
/(60*60);
IF MID_SLEEP_TIME_WKDAY >= 24
  then MID_SLEEP_TIME_WKDAY = MID_SLEEP_TIME_WKDAY - 24;
end;
```

Source variable(s):

SLEA1A_2401. What time do you usually go to bed? (on weekdays)

SLEA2A_2401. What time do you usually wake up? (on weekdays)

17.9 MID_SLEEP_TIME_WKEND (Midpoint sleep time in clock hours; weekend)

This is a numeric variable that calculates the midpoint of self-reported sleep time during weekend. It is in clock hours in decimal point (e.g. 3.5 is 3:30am and 15.75 is 5:45pm). Sleep data was cleaned for creation of midpoint sleep time variables.

```
IF SLEA2C_2401 <= .z OR SLEA1C_2401 <= .z then MID_SLEEP_TIME_WKEND = .;
  ELSE IF SLEA2C_2401 > .z AND SLEA1C_2401 > .z then do;
IF SLEA2C_2401 >= SLEA1C_2401
  then MID_SLEEP_TIME_WKEND = ((SLEA2C_2401 + SLEA1C_2401)/2) / (60*60);
ELSE IF SLEA2C_2401 < SLEA1C_2401
```

```
then MID_SLEEP_TIME_WKEND = ((SLEA2C_2401+86400 + SLEA1C_2401)/2)
/(60*60);
IF MID_SLEEP_TIME_WKEND >= 24
then MID_SLEEP_TIME_WKEND = MID_SLEEP_TIME_WKEND - 24;
end;
```

Source variable(s):

SLEA1C_2401. What time do you usually go to bed? (on weekends)

SLEA2C_2401. What time do you usually wake up? (on weekends)

17.10 WHIIRS (Women's Health Initiative Insomnia Rating Scale)

The score is calculated from non-missing values for any of the following five items: SLEA4, SLEA5, SLEA6, SLEA7 and SLEA11. The original items range from 0 to 4 and the total score from 0 to 20. In HCHS/SOL, SLEA form the scoring for questions 4 to 7 each ranged from 1 to 5 so 1 should be subtracted from each of these items to obtain the score for use in calculating WHIIRS. For example, No, not in the past 4 weeks should give you a score of 0, etc. The scoring for SLEA11 ranges from 0 to 4 and does not need to be modified. Another way to think about it is:

$$\text{WHIIRS} = \text{SLEA4} + \text{SLEA5} + \text{SLEA6} + \text{SLEA7} + \text{SLEA11} - 4$$

Reference:

Levine DW, Kripke DF, Kaplan RM, Lewis MA, Naughton MJ, Bowen DJ, Shumaker SA. Reliability and validity of the Women's Health Initiative Insomnia Rating Scale. *Psychol Assess.* 2003 Jun;15(2):137-48.

Source variable(s):

SLEA4. Did you have trouble falling asleep?

SLEA5. Did you wake up several times at night?

SLEA6. Did you wake up earlier than you planned to?

SLEA7. Did you have trouble getting back to sleep after you woke up too early?

SLEA11. Overall, was your typical night's sleep during the past 4 weeks:

18. DIET

18.1 Diet_Score_JAMA (Diet Score from 1st JAMA paper)

Daviglus ML, Talavera GA, Avilés-Santa ML, et al. Prevalence of major cardiovascular risk factors and cardiovascular diseases among Hispanic/Latino individuals of diverse backgrounds in the United States. JAMA. 2012 Nov 7 ;308(17):1775-84.

Name:	Label:	Definition:
Subscore_SatFat	Subscore for Saturated Fatty Acids	Missing, if PRED_TOTSFA or gender is missing; Else: 5, if .z<PRED_TOTSFA<=20%-tile (gender-specific) for Saturated Fatty Acids intake 4, if 20%-tile< PRED_TOTSFA<=40%-tile; 3, if 40%-tile< PRED_TOTSFA<=60%-tile; 2, if 60%-tile< PRED_TOTSFA<=80%-tile; 1, if 80%-tile< PRED_TOTSFA.
Subscore_Fib	Subscore for Dietary Fiber	Missing, if PRED_FIBER or gender is missing; Else: 1, if .z<PRED_FIBER<=20%-tile (gender-specific) for Dietary Fiber intake 2, if 20%-tile< PRED_FIBER<=40%-tile; 3, if 40%-tile< PRED_FIBER<=60%-tile; 4, if 60%-tile< PRED_FIBER<=80%-tile; 5, if 80%-tile< PRED_FIBER. (the order is reverse compared to Subscore_SatFat)
Subscore_Ca	Subscore for Dietary Calcium	Missing, if pred_CA or gender is missing; Else: 1, if .z<pred_CA<=20%-tile (gender-specific) for Calcium intake 2, if 20%-tile< pred_CA<=40%-tile; 3, if 40%-tile< pred_CA<=60%-tile; 4, if 60%-tile< pred_CA<=80%-tile; 5, if 80%-tile< pred_CA. (the order is same as in Subscore_Fib)
Subscore_K	Subscore for Dietary Potassium	Missing, if pred_K or gender is missing; Else: 1, if .z<pred_K<=20%-tile (gender-specific) for Potassium intake 2, if 20%-tile< pred_K<=40%-tile; 3, if 40%-tile< pred_K<=60%-tile; 4, if 60%-tile< pred_K<=80%-tile; 5, if 80%-tile< pred_K. (the order is same as in Subscore_Fib)
Diet_Score_JAMA	Diet Score from 1st JAMA paper	Missing, if any of Subscore_SatFat, Subscore_Fib, Subscore_Ca, Subscore_K is missing; Else Diet_Score_JAMA=Score_SatFat + Score_Fib + Score_Ca + Score_K.

Source variables (from PRED_NUTR_DERV):

PRED_TOTSFA. Total Saturated Fatty Acids (SFA) (g) - NCI predicted intake (based on DTIA28)

PRED_FIBER. Total dietary fiber (g) - NCI predicted intake (based on DTIA38)

PRED_CA. Calcium (mg) - NCI predicted intake (based on DTIA60)

PRED_K. Potassium (mg) - NCI predicted intake (based on DTIA68)

18.2 Diet_Score_JAMA_C2 (Categorized Diet Score JAMA (1=Below, 2=Above 60th %-tile))

Diet_score_JAMA_c2 =1, if .z<Diet_score_JAMA<60th gender-specific percentile;
Diet_score_JAMA_c2 =2, if Diet_score_JAMA>=60th gender-specific percentile;
Missing, if Diet_Score_JAMA is missing.

Source variable(s):

Diet_Score_JAMA. Diet Score from 1st JAMA paper

18.3 AHEI2010 (Alternative Healthy Eating Index 2010)

The alternative healthy eating index (AHEI-2010) is a measure of diet quality based on foods and nutrients predictive of chronic disease risk (Chiuve SE et al, 2012). It is derived using the following 11 components. The overall score as well as its 11 components (AHEI1 to AHEI11) are available.

- 1) Vegetables without potatoes, servings/day
- 2) Whole Fruit (i.e. does not include fruit juice), servings/day
- 3) Whole grains, servings/day
- 4) Sugar sweetened beverages and fruit juice, servings/day
- 5) Nuts and legumes, servings/day
- 6) Red/processed meat, servings/day
- 7) Trans Fat, % energy
- 8) Long-chain (n-3) fats (EPA+DHA), mg/day
- 9) Polyunsaturated fatty acids (PUFA), % energy;
- 10) Sodium, mg/day;
- 11) Alcohol, drinks/day

AHEI-2010 score is the sum of these 11 individual components' scores, each with a range from 0 (worst) and 10 (best). Hence, **AHEI-2010 takes values from 0 to 110**. Higher scores represent healthy eating habits and lower scores represent unhealthy eating habits. Only scores for components "whole grains", sodium and "alcohol" are gender-specific.

In HCHS/SOL, AHEI-2010 was calculated from available 24hr dietary recall data (one or two recalls per participant) using the NCI method to predict usual intake for each component. A participant has an AHEI-2010 score if he/she has at least one 24-hr dietary recall after exclusions. Recalls with daily energy intake (DTIA20) below the sequence-gender specific 1st percentile or above the 99th percentile or unreliable according to the interviewer (DTIA16) were excluded.

For more details on its derivation please see document "HCHS/SOL Dietary Overview, Methods and Guidelines".

Reference: S E Chiuve et al. (2012) "Alternative Dietary Indices Both Strongly Predict Risk of Chronic Disease"; J. Nutr. 142: 1009 – 1018.

18.4 HEI2010 (Healthy Eating Index 2010)

The Healthy Eating Index-2010 (HEI-2010; Guenther et al., J Acad Nutr Diet 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. The overall score as well as its 12 components (HEI1 to HEI12) are available.

Adequacy

1. Total Fruit
2. Whole Fruit
3. Total Vegetables
4. Greens and Beans
5. Whole Grains
6. Dairy
7. Total Protein Foods
8. Seafood and Plant Proteins
9. Fatty Acids

Moderation

10. Refined Grains
11. Sodium
12. Empty Calories

HEI-2010 score is the sum of these 12 components' scores, that range from 0-5, 0-10, or 0-20. 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 (Guenther et al., J Nutr 2014).

In HCHS/SOL, the HEI-2010 was calculated as the average of available 24hr dietary recalls. A participant has a HEI-2010 score if he/she has at least one 24-hr dietary recall after exclusions. Recalls with daily energy intake (DTIA20) below the sequence-gender specific 1st percentile or above the 99th percentile or unreliable according to the interviewer (DTIA16) were excluded. 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.

For more details on its derivation please see document "HCHS/SOL Dietary Overview, Methods and Guidelines".

References:

- 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.
- 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.

- 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.

18.5 DASH (Version 1 – 2003 Eating Plan)

The dietary approaches to stop hypertension (DASH) diet was initially designed to control blood pressure, and it consists of a diet higher in potassium, calcium, magnesium, fiber, and protein, and lower in total fat, saturated fat, and cholesterol (Karanja et al., 1999). **In HCHS/SOL, we created two versions of the DASH score, both based in Gunther's DASH index and using the average of the two 24hr dietary recalls.**

Version 1. Based on 2003 DASH Eating Plan to match Günther *et al.*, 2009.

Version 2. Based on 2018 DASH Eating Plan which recommends whole grains and fat-free or low-fat dairy products and limits sodium intake.

Günther's DASH index was defined based on the 2003 DASH Eating Plan from the National Institutes of Health, National Heart, Lung, and Blood Institute (NHLBI) (see Box 6, https://www.nhlbi.nih.gov/files/docs/public/heart/hbp_low.pdf). This version has 8 components, each ranging from 0 (worst) to 10 (best). The grains component is the sum of the scores for the total grains and whole grains sub-components, and the dairy component is the sum of the scores for total dairy and low-fat dairy sub-components. Each of the four sub-components ranges from 0 (worst) to 5 (best). **The version 1 of the total DASH score is the sum of all 8 components, and it ranges from 0 to 80.** The overall score as well as all components (DASH1A_2003 to DASH8_2003) are available.

1. Grains:
 - total grains
 - whole grains
2. Vegetables
3. Fruits
4. Dairy:
 - all dairy
 - low-fat dairy
5. Meat/Poultry/Fish/Eggs
6. Nuts/Seeds/Legumes
7. Fats/Oils
8. Sweets

For the scoring of both versions of the DASH, we used as cut-off values the serving sizes for 7 different levels of energy intake (1200/1400/1600/1800/2000/2600/3100) recommended by the NHLBI 2018 DASH Eating Plan. Each participant was assigned an estimated energy requirement (EER) based on sex, age, and physical level activity following the estimated daily calorie needs to maintain calorie balance provided by the Dietary Guidelines for Americans 2015-2020 (Table A2-1 from <https://health.gov/dietaryguidelines/2015/guidelines/appendix-2/>).

For more details on its derivation please see document “HCHS/SOL Dietary Overview, Methods and Guidelines”.

References:

- Günther AL, Liese AD, Bell RA, Dabelea D, Lawrence JM, Rodriguez BL, Standiford DA, Mayer-Davis EJ. Association between the dietary approaches to hypertension diet and hypertension in youth with diabetes mellitus. *Hypertension*. 2009 Jan;53(1):6-12.

- Karanja NM, Obarzanek E, Lin PH, McCullough ML, Phillips KM, Swain JF, Champagne CM, Hoben KP. Descriptive characteristics of the dietary patterns used in the Dietary Approaches to Stop Hypertension Trial. DASH Collaborative Research Group. *J Am Diet Assoc*. 1999 Aug;99(8 Suppl):S19-27.

- *DASH Eating Plan*, NHLBI website. Retrieved from <http://www.nhlbi.nih.gov/health/health-topics/topics/dash/followdash>

18.6 DASH (Version 2 – 2018 Eating Plan)

Because an updated DASH Eating Plan became available in 2018 (<https://www.nhlbi.nih.gov/health-topics/dash-eating-plan>), we created a second version of the DASH score, updating Günther’s DASH index using the most recent recommendations. The 2018 DASH Eating Plan recommends whole grains and fat-free or low-fat dairy products and limits sodium intake. Thus, compared to DASH version 1 we excluded subcomponents total grains and total dairy, and we included sodium as a component. This version of the DASH has 9 components, each ranging from 0 (worst) to 10 (best); **the total DASH score is the sum of the 9 components, and it ranges from 0 to 90.** The overall score as well as all components (DASH1_2018 to DASH9_2018) are available.

1. Whole grains
2. Vegetables
3. Fruits
4. Low-fat dairy
5. Meat/Poultry/Fish/Eggs
6. Nuts/Seeds/Legumes
7. Fats/Oils
8. Sweets
9. Sodium

For more details on its derivation please see document “HCHS/SOL Dietary Overview, Methods and Guidelines”.

18.7 SUPPL_USE (Supplement use in the last 30 days)

This is a 0/1 numeric variable.

If SMIA34>0 then SUPPL_USE = 1 (yes)

If SMIA34=0 then SUPPL_USE = 0 (no)

Source variable:

SMIA34 (NDSR File Supplement recall for the last 30 days)

Variable Modifications

Data release	Variable	# participants changed	Notes
12/2013	DOCTOR_VISIT	5335	Corrected to calculate values for version B for the health care use form (HCE)
12/2013	EDUCATION_C2	295	Categorized missing values based on PIEA21B-PIEA21G, PIEA19B, PIEA21I, and PIEA17
12/2013	EDUCATION_C3	295	Categorized missing values based on PIEA21B-PIEA21G, PIEA19B, PIEA21I, and PIEA17
12/2013	GPAQ_REC_MOD	104	Set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.
12/2013	GPAQ_REC_VIG	106	Set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.
12/2013	GPAQ_REC	106	Set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.
12/2013	GPAQ_WORK_MOD	105	Set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.
12/2013	GPAQ_WORK_VIG	106	Set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.
12/2013	GPAQ_WORK	106	Set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.
12/2013	GPAQ_TOTAL_MOD	106	Set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.
12/2013	GPAQ_TOTAL_VIG	106	Set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.
12/2013	GPAQ_TRSPORT	106	Set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.
12/2013	GPAQ_TOTAL	106	Set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.
12/2013	GPAQ_TOTAL_MET	106	Set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.
12/2013	GPAQ_LEVEL	2516	Corrected a typo in code such that participants who were actually "moderate" were being set to "low". Set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.
12/2013	GPAQ_PAG2008	245	Set participants to missing if they were missing GPAQ_MOD_WEEK and GPAQ_VIG_WEEK. Set participants to missing if they reported more than 1440 total minutes per day of active + sedentary.
12/2013	GPAQ_PAG2008YN	245	Changes reflected from GPAQ_PAG2008
12/2013	INSULIN_FAST	32	Changed missing to .L for .L values of LABA73
5/2014	AGG_PHYS	1	Set to missing due to invalid data for 1 item
5/2014	AGG_MENT	1	Set to missing due to invalid data for 1 item
5/2014	LEFT_ABI	173	ABI definition updated to include occlusion failures
5/2014	RIGHT_ABI	173	ABI definition updated to include occlusion failures
5/2014	MRC_CB	18	Updated to count RSEA3=99 as missing
5/2014	FEV1_FVC_RATIO	122	Set to missing for invalid spirometries
5/2014	AIRFLOW_OBSTRUCTION	122	Set to missing for invalid spirometries
5/2014	POSTBD_PREBD_DIFF	2	Set to missing for invalid spirometries
5/2014	COPD_BY_BD_ALLPRE	335	Set to missing for invalid spirometries
5/2014	CESD10	55	For participants with 1 or 2 missing items, the weighted average of the nonmissing items*10 is used instead of the sum of the nonmissing items
5/2014	STAI10	137	For participants with 1 or 2 missing items, the weighted average of the nonmissing items*10 is used instead of the sum of the nonmissing items
5/2014	MED_ESTROGENS	394	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_ESTROGEN_COMBO	394	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_CONTRACEPTIVES	394	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_PROGESTINS	394	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_SERM	394	Revised criteria for missing values of medication variables and corrected to use MUEA39

Data release	Variable	# participants changed	Notes
5/2014	MED_FERTILITY_REG	394	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_LHRH	394	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_LHRH_AG	394	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_LHRH_ANT	394	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_CURRENT_HRT	394	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_LLD	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_STATIN	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_ASPIRIN	57	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_ANTICOAG_COUMARIN	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_ANTICOAG_HEPARIN	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_ANTIPLATELET	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_CLOPID	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_ANTIDIAB	52	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_ANTICOAG	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_COX2	55	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_OI_STEROID	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_CHEMO	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_NSAID	196	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_METFORMIN	52	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_INSULIN	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_ANTIASTHMATICS	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_ANTIHYPERT	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_CACB	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_ANTIARRHYTHMICS	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_ANTIHYPERT_ACEI	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_CARDIACGLYCOSIDES	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_ANTIANGINAL	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_BB	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_BB_OPTHALMIC	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_ANTIHYPERT_AT2RAS	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_DIURETIC	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_DIURETIC_THIAZIDE	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	MED_FIBARES_NICOACID	51	Revised criteria for missing values of medication variables and corrected to use MUEA39
5/2014	METS_NCEP	2	Revised to use GENDER instead of PIEA1

Data release	Variable	# participants changed	Notes
5/2014	METS_IDF	447	Corrected code to require abdominal obesity, revised to use GENDER instead of PIEA1
5/2014	SASH_LANG	11277	Added item SCEA6, set to missing if more than 1 item is missing
5/2014	SASH_SOC	636	Set to missing if any item is missing
5/2014	ETHIS	93	Set to missing if either item is missing
5/2014	INSULIN_OGTT	2	Corrected to account for .N values of insulin
5/2014	HYPERTENSION	1	Updates due to updated medication variables
5/2014	DIABETES1	2	Updates due to updated medication variables
5/2014	DIABETES_SELF	18	Revised to code missing MHEA16 as no and MHEA16=.Q as missing instead of zero
5/2014	DIABETES3	4	Updates due to updated DIABETES_SELF
5/2014	DIABETES_C4	5	Updates due to updated DIABETES3 and medication variables
5/2014	DIAB_DIAG	9	Updates due to updated DIABETES_SELF and DIABETES3
5/2014	DM_AWARE	14608	Set to missing for nondiabetics
5/2014	DM_CONTROL	13027	Set to missing for nondiabetics
5/2014	IMGEN_C2	94	Recoded 9 to missing
5/2014	EXPOSURE_YEAR	11210	Revised to treat TBEA2=99 or 1 as missing. Corrected code to set never smokers to zero, former and current smokers with negative values to missing and to subtract TBEA2 for former smokers
5/2014	CIGARETTES_YEAR	1323	Updates due to changes in Exposure_Year, set current smokers with TBEA9=0 to missing
5/2014	CIGARETTE_PACK_YEARS	1271	Updates due to changes in Exposure_Year
5/2014	NO_DISABILITY_6ITEM	124	Revised to set participants missing NEE forms to missing
5/2014	HIGH_TOTAL_CHOL	2	Updates due to updated medication variables
5/2014	LEFT_ABI2		Removed, replaced by LEFT_ABI
5/2014	RIGHT_ABI2		Removed, replaced by RIGHT_ABI
5/2014	ABI2_OVERALL		Removed, replaced by ABI_OVERALL
5/2014	ABIGRP2_C4		Removed, replaced by ABIGRP_C4
5/2014	ELEVATED_BP_PLUSMEDS		Removed – intermediate variable
5/2014	HIGH_TRIG		Removed – intermediate variable
5/2014	LOW_HDL		Removed – intermediate variable
5/2014	IFG_NCEP		Removed – intermediate variable
5/2014	ABDOMINAL_OBESITY_NCEP		Removed – intermediate variable
5/2014	HIGH_TRIG_PLUSMEDS		Removed – intermediate variable
5/2014	LOW_HDL_PLUSMEDS		Removed – intermediate variable
5/2014	ABDOMINAL_OBESITY_IDF		Removed – intermediate variable
5/2014	IFG_IDF		Removed – intermediate variable
5/2014	ELEVATED_BP_SELFMEDS		Removed – intermediate variable
5/2014	IFG_NCEP_SELFMEDS		Removed – intermediate variable
5/2014	AIRFLOW_LIMITATION_LLN		Removed – intermediate variable
5/2014	ECG_MAJ_N		Removed – intermediate variable
5/2014	ECG_MIN_N		Removed – intermediate variable
5/2014	IFG		Removed, no scientific value
5/2014	IGT		Removed, no scientific value
5/2014	SASH_ALL		Removed based on SCAS recommendations
12/2019	DM_AWARE	1609	Algorithm correction. Positive cases increase n=479, missing cases decrease n=1212.
12/2019	ESS_GT10	626	Correction on >10 instead >=10 from ESS_GE10, N=626 switched from 1 to 0.
12/2019	HYPERTENSION2_AHA	NA	New baseline derived variable, specs from V2
12/2019	HYPERTENSION_AHA_C5	NA	New baseline derived variable, specs from V2
12/2019	ABIGRP2_C4	NA	Rename from ABIGRP_C4_V2 (V2 was version 2 which is now used to denote visit 2)
12/2019	LEFT_ABI2	NA	Rename from Left_ABI_V2

Data release	Variable	# participants changed	Notes
12/2019	RIGHT_ABI2	NA	Rename from Right_ABI_V2
12/2019	ABI2_OVERALL	NA	Rename from ABI_Overall_V2
12/2019	POSTBD_PREBD_DIFF_GT12_2	NA	Rename from POSTBD_PREBD_DIFF_GT12_V2
12/2019	VALID_SPIROMETRY_POSTBD2	NA	Rename from VALID_SPIROMETRY_POSTBD_V2
12/2019	PRECVD	NA	New baseline derived variable, specs from V2
12/2019	FASTTIME, ELEVATED_BP_PLUSMEDS, ABDOMINAL_OBESITY_IDF, HIGH_TRIG_PLUSMEDS, LOW_HDL_PLUSMEDS, IFG_IDF, IDF_NCEP GFR, IFG_NCEP_SELFMEDS	NA	Keep intermediate variables for Diabetes, Metabolic Syndrome and GFR
12/2019	SUPPL_USE	NA	Added new derived variable
12/2019	EDUCATION_C4	NA	Added new derived variable
12/2019	MID_SLEEP_TIME_WKDAY MID_SLEEP_TIME_WKEND	NA	Added new derived variable
12/2019	HEI1 to HEI12, HEI2010	NA	Added Healthy Eating Index variables
12/2019	DASH2003 and components: DASH1A_2003, DASH1B_2003, DASH2_2003, DASH3_2003, DASH4A_2003, DASH4B_2003, DASH5_2003, DASH6_2003, DASH7_2003, DASH8_2003	NA	Added DASH variable (VERSION 1)
12/2019	DASH2018 and components: DASH1_2018, DASH2_2018, DASH3_2018, DASH4_2018, DASH5_2018, DASH6_2018, DASH7_2018, DASH8_2018, DASH9_2018	NA	Added DASH variable (VERSION 2)
12/2019	SELFMED_ANTIDIAB DIABETES4 DIABETES5	NA	Added per Diabetes SIG decision
12/2019	AGEGROUP_C3 YRSUS_C3	NA	
12/2019	PERIODONTAL_DISEASE	3008	Corrected algorithm
12/2019	PERIO_EXTENT	4	Corrected algorithm
12/2019	METS_NCEP2	66	Revised to use GENDER instead of PIEA1
12/2019	MED_STATIN	79	Two more codes added: 399940 and 409925
12/2019	MED_CACB	34	One more code added: 369930
12/2019	MED_BB	5	One more code added: 369988

APPENDIX

HCHS/SOL DRUG CLASSIFICATION SYSTEM (04/26/2013)

GROUPS 1-16 ANTI-INFECTIVE AGENTS

"010000" = "PENICILLINS"
"011000" = "NATURAL PENICILLINS"
"012000" = "AMINOAMPICILLINS"
"013000" = "PENICILLINASE - RESISTANT PENICILLINS"
"014000" = "EXTENDED - SPECTRUM PENICILLINS"
"015000" = "AMIDINOPENICILLIN"
"019900" = "PENICILLIN COMBINATIONS"
"019905" = "Penicillin-Aminoglycoside Combinations"
"019940" = "PENICILLIN-NSAIA COMBINATIONS"
"020000" = "CEPHALOSPORINS"
"021000" = "CEPHALOSPORINS - 1ST GENERATION"
"022000" = "CEPHALOSPORINS - 2ND GENERATION"
"023000" = "CEPHALOSPORINS - 3RD GENERATION"
"024000" = "CEPHALOSPORINS - 4TH GENERATION"
"029900" = "CEPHALOSPORIN COMBINATIONS"
"030000" = "MACROLIDE ANTIBIOTICS"
"031000" = "ERYTHROMYCINS"
"031099" = "ERYTHROMYCIN COMBINATIONS"
"032000" = "TROLEANDOMYCIN"
"034000" = "AZITHROMYCIN"
"035000" = "CLARITHROMYCIN"
"035200" = "DIRITHROMYCIN"
"035500" = "MIOCAMYCIN"
"035700" = "ROXITHROMYCIN"
"036000" = "SPIRAMYCIN"
"040000" = "TETRACYCLINES"
"049900" = "TETRACYCLINE COMBINATIONS"
"049940" = "TETRACYCLINE-NSAIA COMBINATIONS"
"050000" = "FLUROQUINOLONES"
"059900" = "FLUROQUINOLONE COMBINATIONS"
"059940" = "FLUROQUINOLONE-MACROLIDE COMBINATIONS"
"060000" = "R E S E R V E D"
"070000" = "AMINOGLYCOSIDES"
"080000" = "SULFONAMIDES"
"089900" = "SULFA COMBINATIONS"
"090000" = "ANTIMYCOBACTERIAL AGENTS"
"099900" = "ANTI TB COMBINATIONS"
"100000" = "R E S E R V E D"
"110000" = "ANTIFUNGALS"
"114040" = "IMIDAZOLES"

"114070" = "TRIAZOLES"
 "115000" = "ANTIFUNGAL - GLUCAN SYNTHESIS INHIBITORS (ECHINOCANDINS)"
 "120000" = "ANTIVIRAL"
 "121000" = "ANTIRETROVIRALS"
 "121025" = "ANTIRETROVIRALS - FUSION INHIBITORS"
 "121045" = "ANTIRETROVIRALS - PROTEASE INHIBITORS"
 "121050" = "ANTIRETROVIRALS - RTI-NUCLEOSIDE ANALOGUES-PURINES"
 "121060" = "ANTIRETROVIRALS - RTI-NUCLEOSIDE ANALOGUES-PYRIMIDINES"
 "121080" = "ANTIRETROVIRALS - RTI-NUCLEOSIDE ANALOGUES-THYMIDINES"
 "121085" = "ANTIRETROVIRALS - RTI-NUCLEOTIDE ANALOGUES"
 "121090" = "ANTIRETROVIRALS - RTI-NON-NUCLEOSIDE ANALOGUES"
 "121099" = "ANTIRETROVIRAL COMBINATIONS"
 "122000" = "CMV AGENTS"
 "123500" = "HEPATITIS AGENTS"
 "123520" = "HEPATITIS B AGENTS"
 "123530" = "HEPATITIS C AGENTS"
 "124050" = "HERPES ANTIVIRAL AGENTS - PURINE ANALOGUES"
 "124080" = "HERPES ANTIVIRAL AGENTS - THYMIDINE ANALOGUES"
 "125000" = "INFLUENZA AGENTS"
 "125040" = "NEURAMINIDASE INHIBITORS"
 "126000" = "RESPIRATORY SYNCYTIAL VIRUS (RSV) AGENTS"
 "126040" = "RSV AGENTS - NUCLEOSIDE ANALOGUES"
 "127000" = "MISC. ANTIVIRALS"
 "129900" = "ANTIVIRAL COMBINATIONS"
 "129950" = "NUCLEOSIDE ANALOG & INTERFERON COMBINATIONS"
 "130000" = "ANTIMALARIAL"
 "139900" = "ANTIMALARIAL COMBINATIONS"
 "140000" = "AMEBICIDES"
 "149900" = "AMEBICIDE COMBINATIONS"
 "150000" = "ANTHELMINTIC"
 "159900" = "ANTHELMINTIC COMBINATIONS"
 "160000" = "ANTI-INFECTIVE AGENTS - MISC"
 "161000" = "POLYMYXINS"
 "161500" = "CARBAPENEMS"
 "161599" = "CARBAPENEM COMBINATIONS"
 "162000" = "CHLORAMPHENICOLS"
 "162100" = "KETOLIDES"
 "162200" = "LINCOSAMIDES"
 "162300" = "OXAZOLIDINONES"
 "162500" = "STREPTOGRAMINS"
 "162599" = "STREPTOGRAMIN COMBINATIONS"
 "162700" = "CYCLIC LIPOPEPTIDES"
 "163000" = "LEPROSTATICS"
 "164000" = "ANTIPROTOZOAL AGENTS"
 "164030" = "FOLATE ANTAGONISTS"
 "165000" = "ANTIINFECTIVE ADJUVANTS"

"166000" = "Sepsis Syndrome Agents"
"166040" = "SEPSIS SYNDROME AGENTS - NON-ANTIBIOTIC"
"169900" = "ANTI-INFECTIVE MISC. - COMBINATIONS"

GROUPS 17-20 BIOLOGICALS

"170000" = "VACCINES"
"171000" = "VIRAL VACCINES"
"171099" = "VACCINE COMBINATIONS"
"172000" = "BACTERIAL VACCINES"
"179900" = "MIXED VACCINE COMBINATIONS"
"180000" = "TOXOIDS"
"189900" = "TOXOID COMBINATIONS"
"190000" = "PASSIVE IMMUNIZING AGENTS"
"191000" = "IMMUNE SERUMS"
"192000" = "ANTITOXINS - ANTIVENINS"
"195000" = "MONOCLONAL ANTIBODIES"
"195020" = "ANTIVIRAL MONOCLONAL ANTIBODIES"
"199900" = "PASSIVE IMMUNIZING AGENTS - COMBINATIONS"
"200000" = "BIOLOGICALS MISC"
"201000" = "ALLERGENIC EXTRACTS"
"201099" = "MIXED ALLERGENIC EXTRACTS"

GROUP 21 - ANTINEOPLASTIC AGENTS

"210000" = "ANTINEOPLASTICS"
"211000" = "ALKYLATING AGENTS"
"211010" = "NITROGEN MUSTARDS"
"211020" = "NITROSOUREAS"
"211040" = "IMIDAZOTETRAZINES"
"212000" = "ANTINEOPLASTIC ANTIBIOTICS"
"212500" = "ANTINEOPLASTIC ENZYMES"
"213000" = "ANTIMETABOLITES"
"213300" = "ANTINEOPLASTIC - ANGIOGENESIS INHIBITORS"
"213350" = "VASCULAR ENDOTHELIAL GROWTH FACTOR (VEGF) INHIBITORS"
"213500" = "ANTINEOPLASTIC - ANTIBODIES"
"213530" = "ANTINEOPLASTIC - MONOCLONAL ANTIBODIES"
"213550" = "ANTINEOPLASTIC ANTIBODY-DRUG COMPLEXES"
"213580" = "ANTINEOPLASTIC -ANTIBODY FOR RADIOPHARMACEUTICAL THERAPY"
"214000" = "ANTINEOPLASTIC - HORMONAL AGENTS"
"214020" = "ANDROGENS - ANTINEOPLASTIC"
"214022" = "ANTIADRENALS"
"214024" = "ANTIANDROGENS - ANTINEOPLASTIC"
"214026" = "ANTIESTROGENS - ANTINEOPLASTIC"
"214028" = "AROMATASE INHIBITORS"
"214030" = "ESTROGENS - ANTINEOPLASTIC"

"214035" = "ESTROGEN RECEPTOR ANTAGONIST"
 "214040" = "PROGESTINS - ANTINEOPLASTIC"
 "214050" = "LHRH Analogs"
 "214055" = "GONADOTROPIN RELEASING HORMONE (GNRH) ANTAGONISTS"
 "214500" = "ANTINEOPLASTIC - IMMUNOMODULATORS"
 "215000" = "MIOTIC INHIBITORS"
 "215300" = "ANTINEOPLASTIC ENZYME INHIBITORS"
 "215340" = "ANTINEOPLASTIC - PROTEIN-TYROSINE KINASE INHIBITORS"
 "215360" = "ANTINEOPLASTIC - PROTEASOME INHIBITORS"
 "215500" = "TOPOISOMERASE I INHIBITORS"
 "216000" = "ANTINEOPLASTIC RADIOPHARMACEUTICALS"
 "217000" = "ANTINEOPLASTICS MISC."
 "217030" = "ANTINEOPLASTICS - INTERLEUKINS"
 "217070" = "ANTINEOPLASTICS - PHOTOACTIVATED AGENTS"
 "217080" = "RETINOIDS"
 "217082" = "SELECTIVE RETINOID X RECEPTOR AGONISTS"
 "217500" = "CHEMOTHERAPY RESCUE/ANTIDOTE AGENTS"
 "217540" = "CARDIAC PROTECTIVE AGENTS"
 "217550" = "FOLIC ACID ANTAGONISTS RESCUE AGENTS"
 "217580" = "Urinary Tract Protective Agents"
 "217600" = "CHEMOTHERAPY ADJUNCTS**"
 "217640" = "CHEMOTHERAPY ADJUNCTS - HYPERURICEMIA AGENTS"
 "217650" = "CHEMOTHERAPY ADJUNCTS - KERATINOCYTE GROWTH FACTORS"
 "218000" = "INVESTIGATIONAL ANTINEOPLASTIC"
 "219900" = "ANTINEOPLASTIC COMBINATIONS"
 "219930" = "ANTINEOPLASTIC - ANTIGENIC/IMMUNOLOGIC ADJUVANT COMB"

GROUPS 22-30 ENDOCRINE AND METABOLIC DRUGS

"220000" = "CORTICOSTEROIDS"
 "221000" = "GLUCOCORTICOSTEROIDS"
 "221099" = "STEROID COMBINATIONS"
 "222000" = "MINERALOCORTICIDS"
 "230000" = "ANDROGENS-ANABOLIC"
 "231000" = "ANDROGENS"
 "231099" = "ANDROGEN COMBINATIONS"
 "232000" = "ANABOLIC STEROIDS"
 "232099" = "ANABOLIC STEROID COMBINATIONS"
 "240000" = "ESTROGENS"
 "249900" = "ESTROGEN COMBINATIONS"
 "249910" = "ESTROGEN & ANDROGEN"
 "249915" = "ESTROGEN & ANTIANDROGEN"
 "249920" = "ESTROGEN & ANTIANXIETY AGENT"
 "249930" = "ESTROGEN & PROGESTIN"
 "249940" = "ESTROGEN-ANDROGEN-PROGESTIN"
 "250000" = "CONTRACEPTIVES"

"251000" = "PROGESTIN CONTRACEPTIVES - ORAL"
 "251500" = "PROGESTIN CONTRACEPTIVES - INJECTABLE"
 "252000" = "PROGESTIN CONTRACEPTIVES - IUD"
 "253000" = "PROGESTIN CONTRACEPTIVES - IMPLANTS"
 "254000" = "EMERGENCY CONTRACEPTIVES"
 "259600" = "COMBINATION CONTRACEPTIVES - TRANSDERMAL"
 "259800" = "COMBINATION CONTRACEPTIVES - INJECTABLE"
 "259900" = "COMBINATION CONTRACEPTIVES - ORAL"
 "259910" = "BIPHASIC CONTRACEPTIVES - ORAL"
 "259920" = "TRIPHASIC CONTRACEPTIVES - ORAL"
 "259930" = "EXTENDED-CYCLE CONTRACEPTIVES - ORAL"
 "260000" = "PROGESTINS"
 "270000" = "ANTIDIABETICS"
 "271000" = "INSULIN"
 "271010" = "MIXED INSULIN"
 "271020" = "BEEF INSULIN"
 "271030" = "PORK INSULIN"
 "271040" = "HUMAN INSULIN"
 "271500" = "ANTIDIABETIC - AMYLIN ANALOGS"
 "271700" =
 "272000" = "SULFONYLUREAS"
 "272099" = "SULFOYLUREA COMBINATIONS"
 "272300" = "ANTIDIABETIC - AMINO ACID DERIVATIVES"
 "272340" = "ANTIDIABETIC - D-PHENYLALANINE DERIVATIVES"
 "272500" = "BIGUANIDES"
 "272800" = "MEGLITINIDE ANALOGUES"
 "273000" = "DIABETIC OTHER"
 "273099" = "DIABETIC OTHER - COMBINATIONS"
 "274000" = "ALDOSE REDUCTASE INHIBITORS"
 "275000" = "ALPHA-GLUCOSIDASE INHIBITORS"
 "275500" =
 "276000" = "INSULIN SENSITIZING AGENTS"
 "276070" = "THIAZOLIDINEDIONES"
 "279900" = "ANTIDIABETIC COMBINATIONS"
 "279925" =
 "279970" = "SULFONYLUREA-BIGUANIDE COMBINATIONS"
 "279978" =
 "279990" =
 "280000" = "THYROID AGENTS"
 "281000" = "THYROID HORMONES"
 "281099" = "THYROID COMBINATIONS"
 "283000" = "ANTITHYROID AGENTS"
 "290000" = "OXYTOCICS"
 "292000" = "ABORTIFACIENTS/AGENTS FOR CERVICAL RIPENING"
 "292010" = "ABORTIFACIENTS/CERVICAL RIPENING - PROSTAGLANDINS"
 "292020" = "ANTI-PROGESTATIONAL AGENTS"

"299900" = "OXYTOCIC COMBINATIONS"
 "299910" = "OXYTOCIC MIXTURES"
 "300000" = "ENDOCRINE AND METABOLIC AGENTS - MISC."
 "300200" = "ADRENAL STEROID INHIBITORS"
 "300400" = "CALCIUM REGULATORS"
 "300420" = "BISPHOSPHONATES"
 "300430" = "CALCIUM REGULATOR"
 "300440" = "PARATHYROID HORMONE AND DERIVATIVES"
 "300450" = "CALCIUM REGULATORS - MISC."
 "300500" = "HORMONE RECEPTOR MODULATORS"
 "300530" = "SELECTIVE ESTROGEN RECEPTOR MODULATOR"
 "300600" = "FERTILITY REGULATORS"
 "300620" = "OVULATION STIMULANTS-GONADOTROPINS"
 "360660" = "OVULATION STIMULANTS-SYNTHETIC"
 "300699" = "FERTILITY REGULATOR COMBINATIONS"
 "300700" = "LUTEINIZING HORMONE RELEASING-HORMONES"
 "300800" = "LHRH/GNRH AGONIST ANALOG PITUITARY SUPPRESSANTS"
 "300900" = "GNRH/LHRH ANTAGONISTS"
 "301000" = "GROWTH HORMONE"
 "301500" = "GROWTH HORMONE RELEASING HORMONES"
 "301700" = "SOMATOSTATIC AGENTS"
 "301800" = "GROWTH HORMONE RECEPTOR ANTAGONISTS"
 "302000" = "POSTERIOR PITUITARY HORMONES"
 "302010" = "VASOPRESSIN"
 "303000" = "CORTICOTROPIN"
 "304000" = "PROLACTIN INHIBITORS"
 "304020" = "DOPAMINE RECEPTOR AGONISTS"
 "305000" = "PROGESTERONE RECEPTOR ANTAGONISTS"
 "305020" = "ABORTIFACIENT - PROGESTERONE RECEPTOR ANTAGONISTS"
 "306000" = "MENOPAUSAL SYMPTOMS SUPPRESSANTS"
 "306040" = "HORMONAL AGENTS"
 "308000" = "UTERINE RELAXANTS"
 "309000" = "METABOLIC MODIFIERS"
 "309030" = "CARNITINE REPLENISHER - AGENTS"
 "309036" = "FABRY DISEASE - AGENTS"
 "309040" = "HEREDITARY TYROSINEMIA TYPE 1 (HT-1) TREATMENT - AGENTS"
 "309045" = "HOMOCYSTINURIA TREATMENT - AGENTS"
 "309050" = "HYPERPARATHYROID TREATMENT - VITAMIN D ANALOGS"
 "309052" = "CALCIMIMETIC AGENTS"
 "309065" = "MUCOPOLYSACCHARIDOSIS I (MPS I) - AGENTS"
 "309080" = "UREA CYCLE DISORDER - AGENTS"
 "309900" = "ENDOCRINE AND METABOLIC AGENTS MISC. - COMBINATIONS"

GROUPS 31-40 CARDIOVASCULAR AGENTS

"310000" = "CARDIOTONICS"

"311000" = "PHOSPHODIESTERASE INHIBITORS"
 "312000" = "CARDIAC GLYCOSIDES"
 "318000" = "CARDIOPROTECTANTS"
 "320000" = "ANTIANGINAL AGENTS"
 "321000" = "NITRATES"
 "322000" = "ANTIANGINALS - OTHER"
 "329900" = "ANTIANGINAL COMBINATIONS"
 "329910" = "NITRATE COMBINATIONS"
 "330000" = "BETA BLOCKERS"
 "331000" = "BETA BLOCKERS NON-SELECTIVE"
 "332000" = "BETA BLOCKERS CARDIO-SELECTIVE"
 "333000" = "ALPHA-BETA BLOCKERS"
 "340000" = "CALCIUM CHANNEL BLOCKERS"
 "350000" = "ANTIARRHYTHMICS"
 "350500" = "ANTIARRHYTHMICS TYPE I - NONSPECIFIC"
 "351000" = "ANTIARRHYTHMICS TYPE 1-A"
 "352000" = "ANTIARRHYTHMICS TYPE 1-B"
 "353000" = "ANTIARRHYTHMICS TYPE 1-C"
 "354000" = "ANTIARRHYTHMICS TPYE III"
 "355000" = "MISC. ANTIARRHYTHMIC"
 "360000" = "ANTIHYPERTENSIVE"
 "361000" = "ACE INHIBITORS"
 "361500" = "ANGIOTENSIN II RECEPTOR BLOCKERS"
 "361700" =
 "362000" = "ADRENOLYTIC ANTIHYPERTENSIVES"
 "362010" = "ADRENOLYTICS - CENTRALLY ACTING"
 "362020" = "ADRENOLYTICS - PERIPHERALLY ACTING"
 "362030" = "RESERPINE"
 "362500" = "SELECTIVE ALDOSTERONE RECEPTOR ANTAGONISTS"
 "363000" = "AGENTS FOR PHEOCHROMOCYTOMA"
 "364000" = "VASODILATORS"
 "364010" = "FLUOROQUINOLONE VASODIALATORS"
 "364020" = "DOPAMINE D1 RECEPTOR AGONISTS"
 "365000" = "ANTIHYPERTENSIVE - MAOIS"
 "366000" = "MISC. ANTIHYPERTENSIVES"
 "369900" = "ANTIHYPERTENSIVE COMBINATIONS"
 "369910" = "RESERPINE COMBINATIONS"
 "369915" = "CALCIUM CHANNEL BLOCKER & ACE INHIBITOR COMBINATIONS"
 "369918" = "ACE INHIBITOR & DIURETIC COMBINATIONS"
 "369920" = "BETA BLOCKER DIURETIC COMBINATIONS"
 "369925" = "BETA BLOCKER & CALCIUM CHANNEL BLOCKER COMBINATIONS"
 "369930" =
 "369940" = "ANGIOTENSIN II RECEPTOR BLOCKER & THIAZIDES"
 "369950" = "ADRENOLYTICS-CENTRAL & THIAZIDE COMBINATIONS"
 "369955" = "ADRENOLYTICS-PERIPHERAL & THIAZIDES"
 "369970" = "ANTIHYPERTENSIVES-MAOIS & THIAZIDES"

"369980" = "ANTIHYPERTENSIVES-MISC & THIAZIDES
 "369990" = "VASODILATORS & THIAZIDES"
 "370000" = "DIURETICS"
 "371000" = "CARBONIC ANHYDRASE INHIBITORS"
 "372000" = "LOOP DIURETICS"
 "373000" = "MERCURIAL DIURETICS"
 "374000" = "OSMOTIC DIURETICS"
 "375000" = "POTASSIUM SPARING DIURETICS"
 "376000" = "THIAZIDES AND THIAZIDE-LIKE DIURETICS"
 "379000" = "MISC. DIURETICS"
 "379900" = "COMBINATION DIURETICS"
 "379910" = "DIURETICS & POTASSIUM"
 "379920" = "NON-PRESCRIPTION DIURETICS"
 "389099" = "ANAPHYLAXIS THERAPY AGENTS COMBINATIONS"
 "390000" = "ANTIHYPERLIPIDEMICS"
 "391000" = "BILE SEQUESTRANTS"
 "392000" = "FIBRIC ACID DERIVATIVES"
 "393000" = "INTESTINAL CHOLESTEROL ABSORPTION INHIBITORS"
 "394000" = "HMG-COA REDUCTASE INHIBITORS (STATINS)"
 "394099" = "HMG COA REDUCTASE INHIBITOR COMBINATIONS"
 "394500" = "NICOTINIC ACID DERIVATIVES"
 "395000" = "ANTIHYPERLIPIDEMICS - MISC."
 "399900" = "ANTIHYPERLIPIDEMICS - COMBINATIONS"
 "399920" = "FIBRIC ACID DERIVATIVE COMBINATIONS"
 "399940" = "INTEST CHOLEST ABSORP INHIB-HMG COA REDUCTASE INHIB COMB"
 "400000" = "CARDIOVASCULAR AGENTS - MISC."
 "401000" = "PERIPHERAL VASODILATORS"
 "401099" = "VASODILATOR COMBINATIONS"
 "401500" = "MICROVASODILATORS"
 "401600" = "PULMONARY HYPERTENSION - ENDOTHELIN RECEPTOR ANTAGONISTS"
 "401700" = "PROSTAGLANDIN VASODILATORS"
 "401800" = "VASOACTIVE NATRIURETIC PEPTIDES"
 "402000" = "CARDIOPLEGIC SOLUTIONS"
 "402500" = "VASOCONSTRICTOR INHIBITORS"
 "403000" = "IMPOTENCE AGENTS"
 "403030" = "PROSTAGLANDIN - IMPOTENCE AGENTS"
 "403040" = "SELECTIVE CGMP PHOSPHODIESTERASE TYPE 5 INHIBITORS"
 "403080" = "YOHIMBINE"
 "406000" = "VASOPROTECTANTS"
 "406099" = "VASOPROTECTANT COMBINATIONS"
 "409900" = "CARDIOVASCULAR AGENTS MISC. - COMBINATIONS"
 "409925" = "CALCIUM CHANNEL BLOCKER & HMG COA REDUCTASE INHIBIT COMB"

GROUPS 41-45 RESPIRATORY AGENTS

"410000" = "ANTI HISTAMINES"

"411000" = "ANTIHIISTAMINES - ALKYLAMINES"
 "412000" = "ANTIHIISTAMINES - ETHANOLAMINES"
 "413000" = "ANTIHIISTAMINES - ETHYLENEDIAMINES"
 "414000" = "ANTIHIISTAMINES - PHENOTHIAZINES"
 "415000" = "ANTIHIISTAMINES - PIPERIDINES"
 "415500" = "ANTIHIISTAMINES - NON-SEDATING"
 "416000" = "ANTIHIISTAMINES - MISC."
 "419900" = "ANTIHIISTAMINE COMBINATIONS"
 "419910" = "ANTIHIISTAMINE MIXTURES"
 "419920" = "ANTIHIISTAMINE W/ ANTICHO LINERGICS"
 "419930" = "ANTIHIISTAMINES W/ ANTI ANXIETY AGENTS"
 "420000" = "NASAL AGENTS - SYSTEMIC AND TOPICAL"
 "421000" = "SYMPATHOMIMETIC DECONGESTANTS"
 "421010" = "SYSTEMIC DECONGESTANTS"
 "421020" = "TOPICAL DECONGESTANTS"
 "421030" = "NASAL DECONGESTANT INHALERS"
 "422000" = "NASAL STEROIDS"
 "422500" = "NASAL ANTI-INFECTIVES"
 "422510" = "NASAL ANTIBIOTICS"
 "423000" = "NASAL ANTICHO LINERGICS"
 "424000" = "NASAL ANTIALLERGY"
 "424010" = "NASAL ANTIHIISTAMINES"
 "424050" = "NASAL MAST CELL STABILIZERS"
 "424500" = "NASAL MUCOLYTICS"
 "425000" = "NASAL AGENTS - MISC"
 "425099" = "NASAL AGENTS MISC. - COMBINATIONS"
 "429900" = "NASAL COMBINATIONS"
 "429910" = "DECONGESTANT-ANTIHIISTAMINE"
 "429920" = "DECONGESTANT COMBINATION OTHER"
 "429950" = "DECONGESTANT-NASAL STEROID"
 "429960" = "MAST CELL STABILIZER-ANTIHIISTAMINE"
 "429965" = "MAST CELL STABILIZER-DECONGESTANT-ANALGESIC"
 "430000" = "COUGH/COLD/ALLERGY"
 "431000" = "ANTITUSSIVES"
 "431010" = "ANTITUSSIVE - NARCOTIC"
 "431020" = "ANTITUSSIVE - NONNARCOTIC"
 "432000" = "EXPECTORANTS"
 "432020" = "IODINE EXPECTORANTS"
 "432099" = "EXPECTORANT MIXTURES"
 "433000" = "MUCOLYTICS"
 "434000" = "MISC. RESPIRATORY INHALENTS"
 "434010" = "AROMATIC INHALANTS"
 "439900" = "COUGH/COLD/ALLERGY COMBINATIONS"
 "439910" = "DECONGESTANT - ANALGESIC"
 "439915" = "DECONGESTANT - ANALGESIC - ANTICHO LINERGIC"
 "439920" = "ANTIHIISTAMINE - ANALGESIC"

"439925" = "ANTIHIAMINE - ANALGESIC - ANTICHLINERGIC"
 "439928" = "DECONGESTANT & ANTICHLINERGIC"
 "439930" = "DECONGESTANT & ANTIHIAMINE"
 "439935" = "DECONGEST - ANTIHIAMINE - ANTICHLINERGIC"
 "439940" = "DECONGESTANT - ANTIHIAMINE - ANALGESIC"
 "439945" = "DECONGEST - ANTIHIST - ANALGESIC - ANTICHLIN"
 "439950" = "ANTITUSSIVE COMBOS - NARCOTIC"
 "439951" = "NARCOTIC ANTITUSSIVE - DECONGESTANT"
 "439952" = "NARCOTIC ANTITUSSIVE - ANTIHIAMINE"
 "439953" = "NARCOTIC ANTITUSSIVE - DECONGEST-ANTIHIAMINE"
 "439954" = "NARC ANTIHIAMINE - DECONGEST - ANTIHIAMINE - ANALG"
 "439955" = "NON-NARC ANTIHIAMINE-ANALGESIC"
 "439956" = "NON-NARC ANTIHIAMINE - DECONGESTANT"
 "439957" = "NON-NARC ANTIHIAMINE - ANTIHIAMINE"
 "439958" = "NON-NARC ANTIHIAMINE - DECONGEST - ANTIHIAMINE"
 "439959" = "NON-NARC ANTIHIAMINE - DECONGEST - ANTIHIAMINE - ANALG"
 "439960" = "EXPECTORANT COMBINATIONS"
 "439962" = "DECONGESTANT W/EXPECTORANT"
 "439964" = "ANTIHIAMINE W/EXPECTORANT"
 "439965" = "ANTIHIAMINE-EXPECTORANT-ANALGESIC"
 "439966" = "DECONGESTANT - ANTIHIAMINE W/EXPECTORANT"
 "439967" = "DECONGESTANT-ANALGESIC-EXPECTORANT"
 "439968" = "DECONGEST - ANTIHIAMINE - ANALGESIC W/EXPECTORANT"
 "439969" = "ANTITUSSIVE-DECONGESTANT-ANALGESIC"
 "439970" = "ANTITUSSIVE - EXPECTORANT"
 "439973" = "ANTITUSSIVE - EXPECTORANT - DECONGESTANT"
 "439975" = "ANTITUSSIVE - EXPECTORANT - ANTIHIAMINE"
 "439978" = "ANTITUSSIVE - EXPECTORANT - ANALGESIC"
 "439980" = "ANTITUSSIVE - EXPECTORANT - DECONGEST - ANTIHIAMINE"
 "439983" = "ANTITUSSIVE - EXPECTORANT - DECONGEST - ANALGESIC"
 "439985" = "ANTITUSSIVE - EXPECTORANT - ANTIHIAMINE - ANALGESIC"
 "439988" = "ANTIHIAMINE - EXPECTORANT - DECONGEST - ANTIHIAMINE - ANALG"
 "439989" = "ANTITUSSIVE-ANTIHIAMINE-ANALGESIC"
 "439990" = "MISC. RESPIRATORY COMBINATIONS"
 "440000" = "ANTI-ASTHMATICS AND BRONCHODILATOR AGENTS"
 "441000" = "BRONCHODILATORS - ANTICHLINERGICS"
 "441500" = "ANTI-INFLAMMATORY AGENTS"
 "442000" = "SYMPATHOMIMETICS"
 "442010" = "BETA ADRENERGICS"
 "442020" = "MIXED ADRENERGICS"
 "442099" = "ADRENERGIC COMBINATIONS"
 "443000" = "XANTHINES"
 "444000" = "STEROID INHALANTS"
 "445000" = "LEUKOTRIENE MODULATORS"
 "445040" = "5-LIPOXYGENASE INHIBITORS"
 "445050" = "LEUKOTRIENE RECEPTOR ANTAGONISTS"

"446000" = "ANTIASTHMATIC - MONOCLONAL ANTIBODIES"
 "446030" = "ANTI-IGE MONOCLONAL ANTIBODIES"
 "449900" = "ASTHMA AND BRONCHODILATOR AGENT COMBINATIONS"
 "449905" = "XANTHINE MIXTURES"
 "449910" = "XANTHINE - EXPECTORANTS"
 "449920" = "XANTHINE - SYMPATHOMIMETICS"
 "449922" = "XANTHINE - SYMPATHOMIMETIC - EXPECTORANT"
 "449925" = "XANTHINE - BARBITURATES"
 "449927" = "SYMPATHOMIMETIC - BARBITURATE"
 "449928" = "XANTHINE-ANTIHIISTAMINE-EXPECTORANT"
 "449930" = "XANTHINE - SYMPATHOMIMETIC - BARBITURATE"
 "449932" = "XANTHINE - SYMPATHO - BARBIT - EXPECTOR"
 "449936" = "SYMPATHOMIMETIC-ANTICHOLINERGIC"
 "449938" = "SYMPATHOMIMETIC-ANTITUSSIVE"
 "449940" = "SYMPATHOMIMETIC - EXPECTORANTS"
 "449945" = "SYMPATHOMIMETIC-EXPECTORANT-ANTIHIISTAMINE"
 "449947" = "SYMPATHOMIMETIC-STEROID W/ ANTI-INFECTIVE"
 "449948" = "SYMPATHOMIMETIC W/ ANTI-INFECTIVES"
 "449950" = "XANTHINE - ANTITUSSIVE"
 "449952" = "XANTHINE-ANTITUSSIVE-ANTIHIISTAMINE"
 "449955" = "XANTHINE-ANTITUSSIVE-EXPECTORANT"
 "449980" = "STEROID-ANTIHIISTAMINES"
 "449982" = "STEROID-ANTIHIISTAMINE-EXPECTORANTS"
 "449984" = "STEROID-SYMPATHOMIMETIC-ANTIHIISTAMINE-&/OR EXPECTORANT"
 "449985" = "XANTHINE-ANTIHIISTAMINES"
 "449988" = "Xanthine-Steroids"
 "449990" = "Xanthine-Sympathomimetic-Antihistamine-&/or Expectorant"
 "449993" = "Xanthine-Sympathomimetic-Antitussive-Expectorant"
 "449995" = "Xanthine-Sympath-Antihistamine-Antitussive-Expectorant"
 "450000" = "RESPIRATORY AGENTS - MISC"
 "451000" = "ALPHA - PROTEINASE INHIBITOR (HUMAN)"
 "453000" = "CYSTIC FIBROSIS AGENTS"
 "453040" = "HYDROLYTIC ENZYMES"
 "454000" = "HYPOXIC RESPIRATORY FAILURE AGENTS"
 "455000" = "PLEURAL SCLEROSING AGENTS"

GROUPS 46-52 GASTROINTESTINAL AGENTS

"460000" = "LAXATIVES"
 "461000" = "SALINE LAXATIVES"
 "461099" = "SALINE LAXATIVE MIXTURES"
 "462000" = "STIMULANT LAXATIVES"
 "463000" = "BULK LAXATIVES"
 "464000" = "LUBRICANT LAXATIVES"
 "465000" = "SURFACTANT LAXATIVES"
 "466000" = "MISC. LAXATIVES"

"469900" = "LAXATIVE COMBINATIONS"
 "469910" = "LAXATIVES & DSS"
 "469920" = "BOWEL EVACUATION COMBINATIONS"
 "470000" = "ANTIDIARRHEALS"
 "471000" = "ANTIPERISTALTIC AGENTS"
 "472000" = "GI ADSORBANTS"
 "473000" = "MISC. ANTIDIARRHEAL AGENTS"
 "479900" = "ANTIDIARRHEAL COMBINATIONS"
 "479910" = "DIARRHEA COMBINATIONS-OPIATE"
 "479920" = "DIARRHEA COMBINATIONS-ANTICHOLINERGIC"
 "480000" = "ANTACIDS"
 "481000" = "ANTACIDS - ALUMINUM SALTS"
 "482000" = "ANTACIDS - BICARBONATE"
 "482099" = "ANTACIDS - BICARBONATE COMBINATIONS"
 "483000" = "ANTACIDS - CALCIUM SALTS"
 "484000" = "ANTACIDS - MAGNESIUM SALTS"
 "485000" = "ANTACIDS - SODIUM CITRATE"
 "489900" = "ANTACID COMBINATIONS"
 "489905" = "ANTACID & DIMETHICONE"
 "489910" = "ANTACID & SIMETHICONE"
 "490000" = "ULCER DRUGS"
 "491000" = "ANTISPASMODICS"
 "491010" = "BELLADONNA ALKALOIDS"
 "491020" = "QUATERNARY ANTICHOLINERGICS"
 "491030" = "ANTISPASMODICS"
 "491040" = "ANTICHOLINERGICS"
 "491099" = "ANTICHOLINERGIC COMBINATIONS"
 "492000" = "H-2 ANTAGONISTS"
 "492500" = "ULCER DRUGS - PROSTAGLANDINS"
 "492700" = "PROTON PUMP INHIBITORS"
 "493000" = "MISC. ANTI-ULCER"
 "499900" = "ULCER THERAPY COMBINATIONS"
 "499920" = "ULCER ANTI-INFECTIVE W/ BISMUTH COMBINATIONS"
 "499930" = "ULCER ANTI-INFECTIVE W/ PROTON PUMP INHIBITORS"
 "499935" = "H-2 ANTAGONIST-ANTACID COMBINATIONS"
 "499940" = "H-2 ANTAGONIST-ANTIEMETIC COMBINATIONS"
 "499950" = "H-2 ANTAGONIST-ANTISPASMODIC COMBINATIONS"
 "500000" = "ANTIEMETICS"
 "501000" = "ANTIEMETICS - ANTIDOPAMINERGIC"
 "502000" = "ANTIEMETICS - ANTICHOLINERGIC"
 "502500" = "ANTIEMETICS - 5HT3 ANTAGONISTS"
 "502800" = "SUBSTANCE P/NEUROKININ 1 (NK1) RECEPTOR ANTAGONISTS"
 "503000" = "ANTIEMETICS MISC."
 "503099" = "ANTIEMETICS COMBINATIONS"
 "510000" = "DIGESTIVE AIDS"
 "511000" = "CHOLERETICS"

"511099" = "BILE COMBINATIONS"
 "512000" = "DIGESTIVE ENZYMES"
 "512098" = "DIGESTIVE ENZYMES W/ ANTICHOLINERGICS"
 "512099" = "DIGESTIVE ENZYME COMBINATIONS"
 "513000" = "GASTRIC ACIDIFIERS"
 "513099" = "GASTRIC ACIDIFIERS-COMBINATIONS"
 "514000" = "HYDROCHOLERETICS"
 "519900" = "DIGESTIVE AIDS - MIXTURES"
 "519910" = "DIGESTIVE MIXTURES W/ SIMETHICONE"
 "519920" = "DIGESTIVE MIXTURES W/ ANTICHOLINERGICS"
 "520000" = "GASTROINTESTINAL AGENTS - MISC."
 "521000" = "GALLSTONE SOLUBILIZING AGENTS"
 "521099" = "GALLSTONE SOLUBILIZING AGENTS-COMBINATIONS"
 "521600" = "GASTROINTESTINAL ANTIALLERGY AGENTS"
 "522000" = "ANTIFLATULENTS"
 "522099" = "ANTIFLATULENTS COMBINATIONS"
 "523000" = "GI STIMULANTS"
 "523099" = "GI STIMULANTS COMBINATIONS"
 "524000" = "INTESTINAL ACIDIFIERS"
 "525000" = "INFLAMMATORY BOWEL AGENTS"
 "525050" = "TUMOR NECROSIS FACTOR ALPHA BLOCKERS"
 "525500" = "IRRITABLE BOWEL SYNDROME (IBS) AGENTS"
 "525540" = "IBS AGENT - SELECTIVE 5-HT3 RECEPTOR ANTAGONISTS"
 "525550" = "IBS AGENT - 5-HT4 RECEPTOR PARTIAL AGONISTS"
 "526000" = "HEPATOTROPIC"
 "526099" = "HEPATOTROPIC COMBINATIONS"
 "528000" = "PHOSPHATE BINDER AGENTS"

GROUPS 53-56 GENITOURINARY PRODUCTS

"530000" = "URINARY ANTIINFECTIVES"
 "539900" = "COMBINATION URINARY ANTIINFECTIVES"
 "539905" = "METHENAMINE COMBINATIONS"
 "539910" = "URINARY ANTIINFECTIVE & ANALGESIC"
 "539920" = "URINARY ANTISEPTIC-ANTISPASMODIC &/OR ANALGESICS"
 "540000" = "URINARY ANTISPASMODICS"
 "549900" = "URINARY ANTISPASMODIC COMBINATIONS"
 "550000" = "VAGINAL PRODUCTS"
 "551000" = "VAGINAL ANTIINFECTIVES"
 "551010" = "MISC. VAGINAL ANTIINFECTIVES"
 "551040" = "IMIDAZOLE-RELATED ANTIFUNGALS"
 "551099" = "VAGINAL ANTIINFECTIVE COMBINATIONS"
 "551500" = "VAGINAL ANTIINFLAMMATORY AGENTS"
 "551510" = "VAGINAL CORTICOSTEROIDS"
 "551540" = "VAGINAL NSAIAS"
 "552000" = "DOUCHE PRODUCTS"

"553000" = "SPERMICIDES"
 "553500" = "VAGINAL ESTROGENS"
 "553599" = "VAGINAL ESTROGENS - COMBINATIONS"
 "553700" = "VAGINAL PROGESTINS"
 "554000" = "MISC. VAGINAL PRODUCTS"
 "554110" = "FERTILITY ENHANCERS"
 "554099" = "MISCELLANEOUS VAGINAL COMBINATIONS"
 "560000" = "MISC. GENITOURINARY PRODUCTS"
 "561000" = "ACIDIFIERS"
 "561010" = "PHOSPHATES"
 "561020" = "SYSTEMIC ACIDIFIERS"
 "562000" = "ALKALINIZERS"
 "562020" = "CITRATES"
 "563000" = "URINARY ANALGESICS"
 "564000" = "CYSTINOSIS AGENTS"
 "565000" = "INTERSTITIAL CYSTITIS AGENTS"
 "566000" = "URINARY STONE AGENTS"
 "567000" = "GU IRRIGANTS"
 "567010" = "ANTIINFECTIVE GU IRRIGANTS"
 "568500" = "PROSTATIC HYPERTROPHY AGENTS"
 "568510" = "5-ALPHA REDUCTASE INHIBITORS"
 "568520" = "5-ALPHA REDUCTASE INHIBITORS"

GROUPS 57-60 CENTRAL NERVOUS SYSTEM DRUGS

"570000" = "ANTIANKXIETY AGENTS"
 "571000" = "BENZODIAZEPINES"
 "572000" = "MISC. ANTIANKXIETY AGENTS"
 "580000" = "ANTIDEPRESSANTS"
 "580300" = "ALPHA-2 RECEPTOR ANTAGONISTS (TETRACYCLICS)"
 "581000" = "MAO INHIBITORS"
 "581200" = "MODIFIED CYCLICS"
 "581600" = "SELECTIVE SEROTONIN REUPTAKE INHIBITORS"
 "581800" = "SEROTONIN-NOREPINEPHRINE REUPTAKE INHIBITORS (SNRIS)"
 "582000" = "TRICYCLIC AGENTS"
 "583000" = "MISC. ANTIDEPRESSANTS"
 "590000" = "ANTIPSYCHOTICS"
 "590500" = "BENZAMIDES"
 "590700" = "BENZISOXAZOLES"
 "591000" = "BUTYROPHENONES"
 "591500" = "DIBENZAPINES"
 "591520" = "DIBENZODIAZEPINES"
 "591530" = "DIBENZOTHIAZEPINES"
 "591540" = "DIBENZOXAZEPINES"
 "591570" = "THIENBENZODIAZEPINES"
 "591600" = "DIHYDROINDOLONES"

"591800" = "DIPHENYLBUTYLPYPERIDINES"
"592000" = "PHENOTHIAZINES"
"592500" = "QUINOLINONE DERIVATIVES"
"593000" = "THIOXANTHINES"
"594000" = "MISC. ANTIPSYCHOTICS"
"595000" = "LITHIUM"
"600000" = "HYPNOTICS"
"601000" = "BARBITURATE HYPNOTICS"
"602000" = "NON-BARBITURATE HYPNOTICS"
"602010" = "BENZODIAZEPINE HYPNOTICS"
"602040" = "NON-BENZODIAZEPINE - GABA-RECEPTOR MODULATORS"
"602060" = "SELECTIVE ALPHA2-ADRENORECEPTOR AGONIST SEDATIVES"
"603000" = "ANTI-HISTAMINE HYPNOTICS"
"603099" = "ANTI-HISTAMINE HYPNOTIC COMBINATIONS"
"609900" = "HYPNOTIC COMBINATIONS"

GROUPS 61 STIMULANTS/ANTIOBESITY/ANOREXIANTS

"610000" = "ADHD/ANTI-NARCOLEPSY/ANTI-OBESITY/ANOREXIANTS"
"611000" = "AMPHETAMINES"
"611099" = "AMPHETAMINE MIXTURES"
"612000" = "ANOREXIANTS NON-AMPHETAMINE"
"612099" = "ANOREXIANT COMBINATIONS"
"612500" = "ANTIOBESITY AGENTS"
"612530" = "FAT ABSORPTION DECREASING AGENTS"
"612540" = "MONOAMINE REUPTAKE INHIBITORS"
"612560" = "SEROTONIN REUPTAKE INHIBITORS"
"61300" = "ANALEPTICS"
"613099" = "ANALEPTIC COMBINATIONS"
"613500" = "ATTENTION-DEFICIT/HYPERACTIVITY DISORDER (ADHD) AGENTS"
"613540" = "ADHD AGENT - SELECTIVE NOREPINEPHRINE REUPTAKE INHIBITOR"
"614000" = "STIMULANTS - MISC."

GROUPS 62-63 MISC. PSYCHOTHERAPEUTIC AND NEUROLOGICAL AGENTS

"620000" = "MISC. PSYCHOTHERAPEUTIC AND NEUROLOGICAL AGENTS"
"620500" = "ANTIDEMENTIA AGENTS"
"620510" = "CHOLINOMIMETICS - ACHE INHIBITORS"
"620535" = "N-METHYL-D-ASPARTATE (NMDA) RECEPTOR ANTAGONISTS"
"620540" = "NOOTROPICS"
"621000" = "SMOKING DETERRENTS"
"621099" = "SMOKING DETERRENT COMBINATIONS"
"623800" = "MOVEMENT DISORDER DRUG THERAPY"
"624000" = "MULTIPLE SCLEROSIS AGENTS"
"624030" = "MULTIPLE SCLEROSIS AGENTS - INTERFERONS"
"624050" = "MULTIPLE SCLEROSIS AGENTS - MONOCLONAL ANTIBODIES"

"624500" = "ANTI-CATAPLECTIC AGENTS"
"628000" = "AGENTS FOR CHEMICAL DEPENDENCY"
"628050" = "AGENTS FOR NARCOTIC WITHDRAWAL"
"629900" = "COMBINATION PSYCHOTHERAPEUTICS"
"629920" = "BENZODIAZEPINES & TRICYCLIC AGENTS"
"629940" = "PHENOTHIAZINES & TRICYCLIC AGENTS"
"629950" = "THIENBENZODIAZEPINES & SSRIS"

GROUPS 64-71 ANALGESICS AND ANESTHETICS

"640000" = "ANALGESICS - NONNARCOTIC"
"641000" = "SALICYLATES"
"641099" = "SALICYLATE COMBINATIONS"
"641500" = "ANALGESICS-PEDTIDE CHANNEL BLOCKERS"
"641540" = "SELECTIVE N-TYPE NEURONAL CALCIUM CHANNEL BLOCKERS"
"642000" = "ANALGESICS OTHER"
"642099" = "ANALGESICS - OTHER COMBINATIONS"
"649900" = "ANALGESIC COMBINATIONS"
"649910" = "ANALGESIC-SEDATIVES"
"649920" = "ANALGESIC-ANTICHOLINERGICS"
"650000" = "ANALGESICS - NARCOTIC"
"651000" = "NARCOTIC AGONISTS"
"652000" = "NARCOTIC PARTIAL AGONISTS"
"659900" = "NARCOTIC COMBINATIONS"
"659910" = "CODEINE COMBINATIONS"
"659913" = "DIHYDROCODEINONE COMBINATIONS"
"659915" = "FENTANYL COMBINATIONS"
"659917" = "HYDROCODONE COMBINATIONS"
"659920" = "PROPOXYPHENE COMBINATIONS"
"659930" = "MEPERIDINE COMBINATIONS"
"659940" = "PENTAZOCINE COMBINATIONS"
"659950" = "TRAMADOL COMBINATIONS"
"660000" = "ANALGESICS - ANTI-INFLAMMATORY"
"661000" = "NONSTEROIDAL ANTI-INFLAMMATORY AGENTS (NSAIDS)"
"661005" = "CYCLOOXYGENASE 2 (COX-2) INHIBITORS"
"661010" = "PHENYLBUTAZONES"
"661099" = "NONSTEROIDAL ANTI-INFLAMMATORY AGENT COMBINATIONS"
"662000" = "GOLD COMPOUNDS"
"662500" = "ANTI-RHEUMATIC ANTIMETABOLITE"
"662600" = "INTERLEUKIN-1 RECEPTOR ANTAGONIST (IL-1RA)"
"662700" = "ANTI-TNF-ALPHA - MONOCLONAL ANTIBODIES"
"662800" = "PYRIMIDINE SYNTHESIS INHIBITORS"
"662900" = "SOLUBLE TUMOR NECROSIS FACTOR RECEPTOR AGENTS"
"663000" = "MISC. ANTI-RHEUMATIC"
"663099" = "MISC. ANTI-RHEUMATIC COMBINATIONS"
"670000" = "MIGRAINE PRODUCTS"

"673000" = "CARBOXYLIC ACID DERIVATIVES"
"674000" = "SEROTONIN AGONISTS"
"674060" = "SELECTIVE SEROTONIN AGONISTS 5-HT(1)"
"679900" = "MIGRAINE COMBINATION"
"679910" = "ERGOT COMBINATIONS"
"680000" = "GOUT AGENTS"
"681000" = "URICOSURICS"
"689900" = "COMBINATION GOUT DRUGS"
"690000" = "LOCAL ANESTHETICS - PARENTERAL"
"691000" = "LOCAL ANESTHETICS - AMIDES"
"692000" = "LOCAL ANESTHETICS - ESTERS"
"699900" = "LOCAL ANESTHETIC COMBINATIONS"
"699910" = "LOCAL ANESTHETIC & SYMPATHOMIMETIC"
"700000" = "GENERAL ANESTHETICS"
"700500" = "ANESTHETIC GASSES"
"701000" = "BARBITURATE ANESTHETICS"
"702000" = "VOLATLE ANESTHETICS"
"704000" = "MISC. ANESTHETICS"
"704099" = "ANESTHETIC COMBINATIONS"

GROUPS 72-76 NEUROMUSCULAR DRUGS

"720000" = "ANTICONVULSANT"
"721000" = "ANTICONVULSANTS - BENZODIAZEPINES"
"721200" = "CARBAMATES"
"721700" = "GABA MODULATORS"
"722000" = "HYDANTOINS"
"723000" = "OXAZOLIDINEDIONES"
"724000" = "SUCCINIMIDES"
"725000" = "VALPROIC ACID"
"726000" = "MISC. ANTICONVULSANTS"
"726099" = "ANTICONVULSANT COMBINATIONS"
"730000" = "ANTIPARKINSONIAN AGENTS"
"731000" = "ANTIPARKINSONIAN ANTICHOLINERGICS"
"731500" = "ANTIPARKINSON COMT INHIBITORS"
"731520" = "CENTRAL/PERIPHERAL COMT INHIBITORS"
"731530" = "PERIPHERAL COMT INHIBITORS"
"732000" = "ANTIPARKINSONIAN DOPAMINERGIC"
"732030" = "D2 CLASS DOPAMINE RECEPTOR AGONISTS"
"732099" = "LEVODOPA COMBINATIONS"
"733000" = "ANTIPARKINSONIAN MONOAMINE OXIDASE INHIBITOR"
"734000" = "ANTIPARKINSON ADJUVANTS"
"734030" = "DECARBOXYLASE INHIBITORS"
"740000" = "NEUROMUSCULAR BLOCKERS"
"741000" = "DEPLOORIZING MUSCLE RELAXANTS"
"742000" = "NONDEPLOORIZING MUSCLE RELAXANTS"

"744000" = "NEUROMUSCULAR BLOCKING AGENT - NEUROTOXINS"
 "745000" = "ALS AGENTS"
 "745030" = "BENZATHIAZOLES"
 "750000" = "MUSCULOSKELETAL MUSCLE RELAXANTS"
 "751000" = "CENTRAL MUSCLE RELAXANTS"
 "752000" = "DIRECT MUSCLE RELAXANTS"
 "753000" = "MISC. MUSCLE RELAXANTS"
 "758000" = "VISCOSUPPLEMENTS"
 "758400" = "ARTICULAR CARTILAGE REPAIR THERAPY"
 "759900" = "MUSCLE RELAXANT COMBINATIONS"
 "759930" = "MUSCLE RELAXANT-STEROID COMBINATIONS"
 "760000" = "ANTIMYASTHENIC AGENTS"
 "769900" = "ANTIMYASTHENIC COMBINATIONS"

GROUPS 77-81 NUTRITIONAL PRODUCTS

"770000" = "VITAMINS"
 "771000" = "WATER SOLUBLE VITAMINS"
 "771010" = "VITAMIN B-1"
 "771020" = "VITAMIN B-2"
 "771030" = "VITAMIN B-3"
 "771040" = "VITAMIN B-5"
 "771050" = "VITAMIN B-6"
 "771060" = "BIOTIN"
 "771070" = "PABA"
 "771080" = "VITAMIN C"
 "772000" = "OIL SOLUBLE VITAMINS"
 "772010" = "VITAMIN A"
 "772020" = "VITAMIN D"
 "772030" = "VITAMIN E"
 "772040" = "VITAMIN K"
 "773000" = "MISC. NUTRITIONAL FACTORS"
 "773030" = "BIOFLAVINOIDS"
 "773099" = "BIOFLAVONOID COMBINATIONS"
 "780000" = "MULTIVITAMINS"
 "781000" = "VITAMIN MIXTURES"
 "781010" = "VITAMINS A & D"
 "781015" = "VITAMINS A & D W/ C"
 "781017" = "VITAMINS A, C, D & E"
 "781020" = "VITAMINS ACE & ZN"
 "781030" = "VITAMINS B 1-2-3"
 "781040" = "VITAMINS C & E"
 "781044" = "NIACIN W/ INOSITOL"
 "781045" = "NIACIN W/ C"
 "781046" = "NIACINAMIDE W/ C"
 "781047" = "NIACINAMIDE W/ E"

"781048" = "NIACINAMIDE W/ ZINC & FOLIC ACID"
 "781049" = "NIACINAMIDE W/ ZINC-COPPER & FOLIC ACID"
 "781050" = "VITAMINS B1 & B6"
 "781060" = "VITAMINS B1, B6 & B12"
 "781100" = "B-COMPLEX VITAMINS"
 "781110" = "BREWERS YEAST"
 "781200" = "B-COMPLEX W/ C"
 "781205" = "B-COMPLEX W/ C & MG"
 "781210" = "B-COMPLEX W/ C + MG ZN"
 "781220" = "B-COMPLEX W/ C & E"
 "781225" = "B-COMPLEX W/ C & E + ZN"
 "781230" = "B-COMPLEX W/ C & CALCIUM"
 "781300" = "B-COMPLEX W/ FOLIC ACID"
 "781320" = "B-COMPLEX W/BIOTIN & FOLIC ACID"
 "781330" = "B-COMPLEX W/ C FOLIC ACID"
 "781334" = "B-COMPLEX W/ E & FOLIC ACID"
 "781340" = "B-COMPLEX W/ IRON & FOLIC ACID"
 "781350" = "B-COMPLEX W/ C-MIN-FE & FOLIC ACID"
 "781355" = "B-COMPLEX W/ C-BIOTIN-FE & FOLIC ACID"
 "781360" = "B-COMPLEX W/ C-ZN & FOLIC ACID"
 "781365" = "B-COMPLEX W/ LYSINE-ZN & FOLIC ACID"
 "781367" = "B-COMPLEX W/ LYSINE-MIN-FE & FOLIC ACID"
 "781370" = "B-COMPLEX W/ C-BIOTIN-E-MINERALS & FOLIC ACID"
 "781375" = "B-COMPLEX W/ C-BIOTIN-MINERALS & FOLIC ACID"
 "781400" = "B-COMPLEX W/ IRON"
 "781500" = "B-COMPLEX W/ MINERALS"
 "781600" = "BIOFLAVONOID PRODUCTS"
 "781700" = "BIOTIN W/ VITAMIN C"
 "782000" = "MULTIVITAMINS"
 "782010" = "HEXAVITAMINS"
 "782020" = "MULTIPLE VITAMIN & APPETITE STIMULANT"
 "782100" = "MULTIPLE VITAMINS W/ IRON"
 "783100" = "MULTIPLE VITAMINS W/ MINERALS"
 "783400" = "MULTIPLE VITAMINS W/ FLUORIDE"
 "783500" = "MULTIPLE VITAMINS W/ CALCIUM"
 "783600" = "MULTIPLE VITAMINS W/ MINERALS & CALCIUM-FOLIC ACID"
 "784000" = "PEDIATRIC VITAMINS"
 "784015" = "PEDIATRIC VITAMINS A & D W/ C"
 "784100" = "PEDIATRIC MULTIPLE VITAMINS"
 "784200" = "PED MULTIPLE VITAMINS W/ MINERALS"
 "784300" = "PED MV W/ IRON"
 "784400" = "PED MV W/ FLUORIDE"
 "784405" = "PED VITAMINS ACD W/FLUORIDE"
 "784410" = "PED MV W/FLUORIDE"
 "784500" = "PED MULTIPLE VITAMINS W/FL & FE"
 "784520" = "PED VITAMINS ACD FLUORIDE & IRON"

"785000" = "SPECIALTY VITAMINS PRODUCTS"
 "785100" = "PRENATAL VITAMINS"
 "785110" = "PRENATAL MV & MINERALS W/ IRON"
 "785120" = "PRENATAL MV & MINERALS W/ IRON & FA"
 "785130" = "PRENATAL MV & MINERALS W/ FA"
 "785140" = "PRENATAL MV & MIN W/FE-FA-CA"
 "785150" = "PRENATAL MV & MIN W/FE-FA-CA-OMEGA 3 FISH OIL"
 "785200" = "VITAMINS W/ LIPOTROPICS"
 "785300" = "VITAMINS W/ HORMONES"
 "786000" = "HEMATINIC-VITAMIN PRODUCTS"
 "786100" = "IRON W/ VITAMINS"
 "786200" = "B-12 W/ VITAMINS"
 "786300" = "IRON & B12 W/ VITAMINS"
 "790000" = "MINERALS & ELECTROLYTES"
 "790500" = "BICARBONATES"
 "791000" = "CALCIUM"
 "791099" = "CALCIUM COMBINATIONS"
 "792000" = "CHLORIDE"
 "793000" = "FLUORIDE"
 "793099" = "FLUORIDE COMBINATIONS"
 "793500" = "IODINE PRODUCTS"
 "793599" = "IODINE COMBINATIONS"
 "794000" = "MAGNESIUM"
 "794099" = "MAGNESIUM COMBINATIONS"
 "795000" = "MANGANESE"
 "795099" = "MANGANESE COMBINATIONS"
 "796000" = "PHOSPHATE"
 "797000" = "POTASSIUM"
 "797099" = "POTASSIUM COMBINATIONS"
 "797500" = "SODIUM"
 "798000" = "ZINC"
 "798099" = "ZINC COMBINATIONS"
 "798500" = "MINERAL COMBINATIONS"
 "799000" = "TRACE MINERALS"
 "799099" = "TRACE MINERAL COMBINATIONS"
 "799900" = "ELECTROLYTE MIXTURES"
 "799910" = "ELECTROLYTES ORAL"
 "799920" = "ELECTROLYTES PARENTERAL"
 "799930" = "ELECTROLYTES & DEXTROSE"
 "799940" = "ELECTROLYTES & INVERT SUGAR"
 "799950" = "PARENTERAL ELECTROLYTES W/ FRUCTOSE"
 "800000" = "NUTRIENTS"
 "801000" = "CARBOHYDRATE"
 "802000" = "LIPIDS"
 "802030" = "FATTY ACIDS"
 "802099" = "LIPID COMBINATIONS"

"803000" = "PROTEIN"
 "803010" = "PROTEIN PRODUCTS"
 "803020" = "AMINO ACID MIXTURES"
 "803030" = "AMINO ACIDS-SINGLE"
 "803099" = "PROTEIN COMBINATIONS"
 "804000" = "LIPOTROPICS"
 "804099" = "LIPOTROPIC COMBINATIONS"
 "805000" = "MISC. NUTRITIONAL SUBSTANCES"
 "805099" = "MISC. NUTRITIONAL SUBSTANCES COMBINATIONS"
 "810000" = "DIETARY PRODUCTS"
 "811000" = "INFANT FOODS"
 "812000" = "NUTRITIONAL SUPPLEMENTS"
 "812005" = "NUTRITIONAL SUPPLEMENTS W/ DYE"
 "812010" = "NUTRITIONAL SUPPLEMENTS - DIET AIDS"
 "813000" = "TUBE FEEDINGS"
 "814000" = "NUTRITIONAL SUBSTITUTES"
 "814010" = "SALT SUBSTITUTES"
 "814020" = "SWEETNERS"
 "819000" = "NUTRITIONAL MODIFIERS"

GROUPS 82-85 HEMATOLOGICAL AGENTS

"820000" = "HEMATOPOETIC AGENTS"
 "821000" = "COBALAMINS"
 "821010" = "LIVER PREPARATIONS"
 "821500" = "INTRINSIC FACTOR"
 "822000" = "FOLIC ACID/FOLATES"
 "823000" = "IRON"
 "824000" = "HEMATOPOIETIC GROWTH FACTORS"
 "824010" = "ERYTHROPOIETINS"
 "824015" = "GRANULOCYTE COLONY-STIMULATING FACTORS (G-CSF)"
 "824020" = "GRANULOCYTE/MACROPHAGE COLONY-STIMULATING FACTOR(GM-CSF)"
 "824030" = "INTERLEUKINS"
 "824040" = "STEM CELL FACTORS (SCF)"
 "827000" = "AGENTS FOR GAUCHER DISEASE"
 "828000" = "AGENTS FOR SICKLE CELL ANEMIA"
 "828030" = "CYTOTOXIC AGENTS"
 "829900" = "HEMOATOPOETIC MIXTURES"
 "829910" = "COBALAMIN COMBINATIONS"
 "829915" = "FOLIC ACID/FOLATE COMBINATIONS"
 "829920" = "IRON COMBINATIONS"
 "829930" = "IRON W/ B12"
 "829940" = "IRON W/ FOLIC ACID"
 "829950" = "IRON-B12-FOLATE"
 "829970" = "LIVER EXTRACT COMBINATIONS"
 "830000" = "ANTICOAGULANTS"

"831000" = "HEPARINS"
 "831010" = "LOW MOLECULAR WEIGHT HEPARINS"
 "831030" = "SYNTHETIC HEPARINOID-LIKE AGENTS"
 "832000" = "COUMARIN ANTICOAGULANTS"
 "833000" = "INDANDIONE ANTICOAGULANTS"
 "833300" = "THROMBIN INHIBITORS"
 "833340" = "THROMBIN INHIBITORS - HIRUDIN TYPE"
 "833370" = "THROMBIN INHIBITORS - SELECTIVE DIRECT & REVERSIBLE"
 "834000" = "IN VIRO ANTICOAGULANTS"
 "834099" = "IN VITRO ANTICOAGULANT COMBINATIONS"
 "840000" = "HEMOSTATICS"
 "841000" = "HEMOSTATICS - SYSTEMIC"
 "841099" = "SYSTEMIC HEMOSTATIC COMBINATIONS"
 "842000" = "HEMOSTATICS - TOPICAL"
 "842099" = "HEMOSTATIC COMBINATIONS - TOPICAL"
 "850000" = "MISC. HEMATOLOGICAL AGENTS"
 "851000" = "ANTIHEMOPHILIC PRODUCTS"
 "851500" = "PLATELET AGGREGATION INHIBITORS"
 "851530" = "GLYCOPROTEIN IIB/IIIA RECEPTOR INHIBITORS"
 "851550" = "MONOCLONAL AGENTS"
 "851555" = "PHOSPHODIESTERASE III INHIBITORS"
 "851560" = "QUINAZOLINE AGENTS"
 "851580" = "THIENOPYRIDINE DERIVATIVES"
 "851599" = "PLATELET AGGREGATION INHIBITOR COMBINATIONS"
 "852000" = "HEMATORHEOLOGICAL"
 "852500" = "HEMIN"
 "852700" = "IN VITRO HEMATOLOGIC AGENTS"
 "852770" = "RED CELL WASHING AGENTS"
 "852775" = "RED CELL PRESERVATION AGENTS"
 "822780" = "RED CELL REJUVENATION AGENTS"
 "853000" = "PLASMA EXPANDERS"
 "854000" = "PLASMA PROTEINS"
 "855000" = "PROTAMINE"
 "856000" = "THROMBOLYTIC ENZYMES"
 "856010" = "TISSUE PLASMINOGEN ACTIVATOR"
 "857000" = "HEMATOLOGIC OXYGEN TRANSPORTER"

GROUPS 86-91 TOPICAL PRODUCTS

"860000" = "OPHTHALMIC AGENTS"
 "861000" = "OPHTHALMIC ANTIINFECTIVES"
 "861010" = "OPHTHALMIC ANTIBIOTICS"
 "861020" = "OPHTHALMIC SULFONAMIDES"
 "861030" = "OPHTHALMIC ANTIVIRALS"
 "861040" = "OPHTHALMIC ANTIFUNGAL"
 "861050" = "OPHTHALMIC ANTISEPTICS"

"861099" = "OPHTHALMIC ANTIINFECTIVE COMBINATIONS"
 "862000" = "ARTIFICIAL TEARS AND LUBRICANTS"
 "862010" = "ARTIFICIAL TEAR SOLUTIONS"
 "862020" = "ARTIFICIAL TEAR OINTMENTS"
 "862025" = "ARTIFICIAL TEAR GELS"
 "862030" = "ARTIFICIAL TEAR INSERT"
 "862040" = "GONIOSCOPIIC SOLUTION"
 "862099" = "ARTIFICIAL TEAR AND LUBRICANT COMBINATIONS"
 "862500" = "BETA-BLOCKERS - OPHTHALMIC"
 "862599" = "BETA-BLOCKERS - OPHTHALMIC COMBINATIONS"
 "863000" = "OPHTHALMIC STEROIDS"
 "863099" = "OPHTHALMIC STEROID COMBINATIONS"
 "863300" = "PROSTAGLANDINS - OPHTHALMIC"
 "863500" = "CYCLOPLEGIC MYDRIATICS"
 "863599" = "CYCLOPLEGIC MYDRIATICS COMBINATIONS"
 "864000" = "OPHTHALMIC DECONGESTANTS"
 "864099" = "OPHTHALMIC DECONGESTANT COMBINATIONS"
 "865000" = "MIOTICS"
 "865010" = "MIOTICS - DIRECT ACTING"
 "865020" = "MIOTICS - CHOLINESTERASE INHIBITORS"
 "865099" = "MIOTIC COMBINATIONS"
 "866000" = "OPHTHALMIC ADRENERGIC AGENTS"
 "866020" = "OPHTHALMIC SELECTIVE ALPHA ADRENERGIC AGONISTS"
 "866500" = "OPHTHALMIC - ANGIOGENESIS INHIBITORS"
 "866550" = "SELECTIVE VASCULAR ENDOTHELIAL GROWTH FACTOR ANTAGONIST"
 "867000" = "OPHTHALMIC PHOTODYNAMIC THERAPY AGENTS"
 "867200" = "OPHTHALMIC IMMUNOMODULATORS"
 "867500" = "OPHTHALMIC LOCAL ANESTHETICS"
 "867599" = "OPHTHALMIC LOCAL ANESTHETIC - COMBINATIONS"
 "867800" = "OPHTHALMIC SURGICAL AIDS"
 "867899" = "OPHTHALMIC SURGICAL AIDS - COMBINATIONS"
 "868000" = "MISC. OPHTHALMICS"
 "868010" = "OPHTHALMIC ENZYMES"
 "868020" = "OPHTHALMIC ANTIALLERGIC"
 "868023" = "OPHTHALMIC CARBONIC ANHYDRASE INHIBITORS"
 "868030" = "OPHTHALMIC IRRIGATION SOLUTIONS"
 "868040" = "OPHTHALMIC HYPEROSMOLAR PRODUCTS"
 "868050" = "OPHTHALMIC NSAIA'S AGENT"
 "868060" = "OPHTHALMIC DIAGNOSTIC PRODUCTS"
 "868070" = "OPHTHALMICS MISC. - OTHER"
 "868099" = "MISC. OPHTHALMIC COMBINATIONS"
 "869000" = "CONTACT LENS SOLUTIONS"
 "869010" = "HARD LENS PRODUCTS"
 "869020" = "SOFT LENS PRODUCTS"
 "869030" = "OXYGEN PERMEABLE LENS PRODUCTS"
 "869050" = "HARD/SOFT/GAS PERMEABLE PRODUCTS"

"870000" = "OTIC AGENTS"
 "871000" = "OTIC ANTINFECTIVES"
 "871099" = "OTIC ANTINFECTIVES COMBINATIONS"
 "872000" = "OTIC ANALGESICS"
 "873000" = "OTIC STEROIDS"
 "874000" = "OTIC AGENTS MISC."
 "877000" = "OTIC AGENTS - FOR EXTERNAL EAR"
 "877097" = "OTIC - EXTERNAL ANALGESIC COMBINATIONS"
 "879900" = "OTIC COMBINATIONS"
 "879910" = "OTIC STEROID-ANTI-INFECTIVE COMBINATIONS"
 "879920" = "OTIC ANALGESIC COMBINATIONS"
 "879930" = "OTIC ANTIFUNGAL COMBINATIONS"
 "880000" = "MOUTH/THROAT/DENTAL AGENTS"
 "881000" = "ANTIINFECTIVES - THROAT"
 "881099" = "ANTI-INFECTIVE COMBINATIONS - THROAT"
 "881500" = "ANTISEPTICS - MOUNT/THROAT"
 "881599" = "ANTISEPTIC COMBINATIONS - MOUTH/THROAT"
 "882000" = "LOZENGES"
 "882099" = "LOZENGE - COMBINATIONS"
 "882500" = "STEROIDS - MOUTH/THROAT"
 "882599" = "STEROID COMBINATIONS - MOUTH/THROAT"
 "883000" = "MOUTHWASHES"
 "883099" = "MOUTHWASH COMBINATIONS"
 "883500" = "ANESTHETICS, TOPICAL ORAL"
 "883599" = "ANESTHETICS, TOPICAL ORAL - COMBOS"
 "884000" = "DENTAL PRODUCTS"
 "884003" = "DENTAL ASTRINGENTS"
 "884005" = "DENTAL DISINFECTANTS"
 "884007" = "DENTAL HEMOSTATICS"
 "884010" = "FLUORIDE DENTAL RINSE"
 "884020" = "FLUORIDE DENTAL PRODUCTS"
 "884030" = "FLUORIDE TOOTHPASTE"
 "884070" = "DENTAL AIDS"
 "884099" = "DENTAL PRODUCTS - COMBINATIONS"
 "885000" = "MISC. THROAT PRODUCTS"
 "885010" = "ARTIFICIAL SALIVA"
 "885015" = "SALIVA STIMULANTS"
 "885020" = "PROTECTANTS"
 "890000" = "ANORECTAL"
 "891000" = "RECTAL STEROIDS"
 "891500" = "INTRARECTAL STEROIDS"
 "892000" = "RECTAL LOCAL ANESTHETICS"
 "893000" = "MISC. RECTAL PRODUCTS"
 "894000" = "RECTAL PROTECTANTS - EMOLLIENTS"
 "899900" = "RECTAL COMBINATIONS"
 "899910" = "RECTAL ANESTHETIC/STEROIDS"

"899920" = "RECTAL ANESTHETIC COMBINATIONS"
 "899930" = "RECTAL STEROID COMBINATIONS"
 "899940" = "MISC. RECTAL COMBINATIONS"
 "900000" = "DERMATOLOGICALS"
 "900500" = "ACNE PRODUCTS"
 "900510" = "ACNE ANTIBIOTICS"
 "900520" = "ACNE CLEANSERS"
 "900599" = "ACNE COMBINATIONS"
 "900600" = "ROSACEA AGENTS"
 "900700" = "ANALGESICS - TOPICAL"
 "900799" = "ANALGESIC COMBINATIONS - TOPICAL"
 "901000" = "ANTIBIOTICS - TOPICAL"
 "901098" = "ANTIBIOTIC MIXTURES TOPICAL"
 "901099" = "ANTIBIOTIC STEROID COMBINATIONS - TOPICAL"
 "901500" = "ANTIFUNGALS - TOPICAL"
 "901540" = "IMIDAZOLE-RELATED ANTIFUNGALS - TOPICAL"
 "901599" = "ANTIFUNGALS - TOPICAL COMBINATIONS"
 "902000" = "ANTI HISTAMINES-TOPICAL"
 "902099" = "ANTI HISTAMINES - TOPICAL COMBINATIONS"
 "902100" = "ANTI INFLAMMATORY AGENTS"
 "902199" = "ANTI-INFLAMMATORY COMBINATIONS - TOPICAL"
 "902200" = "ANTI PRURITICS"
 "902299" = "ANTI PRURITIC COMBINATIONS - TOPICAL"
 "902500" = "ANTIPSORATICS"
 "902505" = "ANTIPSORIATICS - SYSTEMIC"
 "902510" = "ANTIPSORATIC, TAR CONTAINING"
 "902599" = "ANTIPSORATIC COMBINATIONS"
 "903000" = "ANTISEBORRHEIC PRODUCTS"
 "903099" = "ANTISEBORRHEIC COMBINATIONS"
 "903500" = "ANTIVIRAL - TOPICAL"
 "903599" = "ANTIVIRAL TOPICAL COMBINATIONS"
 "903700" = "ANTINEOPLASTIC OR PREMALIGNANT LESION AGENTS - TOPICAL"
 "903720" = "ANTINEOPLASTIC ANTIMETABOLITES - TOPICAL"
 "903740" = "ANTINEOPLASTIC OR PREMALIGNANT LESIONS - TOPICAL NSAID'S"
 "903750" = "PHOTODYNAMIC THERAPY AGENTS - TOPICAL"
 "903760" = "ANTINEOPLASTIC RETINOIDS - TOPICAL"
 "903762" = "TOPICAL SELECTIVE RETINOID X RECEPTOR AGONISTS"
 "904000" = "BATH PRODUCTS"
 "904500" = "BURN PRODUCTS"
 "905000" = "CAUTERIZING AGENTS"
 "905099" = "CAUTERIZING AGENT COMBINATIONS"
 "905200" = "TAR PRODUCTS"
 "905299" = "TAR COMBINATIONS"
 "905500" = "CORTICOSTEROIDS - TOPICAL"
 "905598" = "STEROID-LOCAL ANESTHETICS"
 "905599" = "TOPICAL STEROID COMBINATIONS"

"906000" = "DIAPER RASH PRODUCTS"
 "906500" = "EMOLLIENTS"
 "906599" = "EMOLLIENT COMBINATIONS"
 "906600" = "EMOLLIENT/KERATOLYTIC AGENTS"
 "907000" = "ENZYMES - TOPICAL"
 "907099" = "ENZYME MIXTURES - TOPICAL"
 "907300" = "HAIR GROWTH AGENTS"
 "907360" = "TYPE II 5-ALPHA REDUCTASE INHIBITORS"
 "907380" = "VASCULAR AGENTS"
 "907400" = "HAIR REDUCTION AGENTS"
 "907450" = "ORNITHINE DECARBOXYLASE (ODC) INHIBITORS - TOPICAL"
 "907485" = "SURGICAL DEPILATORIES"
 "907500" = "KERATOLYTIC/ANTIMITOTIC AGENTS"
 "907599" = "KERATOLYTIC AND/OR ANTIMITOTIC AGENTS COMBINATIONS"
 "907700" = "IMMUNOMODULATING AGENTS - TOPICAL"
 "907730" = "IMMUNOMODULATORS IMIDAZOQUINOLINAMINES - TOPICAL"
 "907800" = "IMMUNOSUPPRESSIVE AGENTS - TOPICAL"
 "907840" = "MACROLIDE IMMUNOSUPPRESSANTS - TOPICAL"
 "908000" = "LINIMENTS"
 "908099" = "LINIMENT COMBINATIONS"
 "908500" = "LOCAL ANESTHETICS - TOPICAL"
 "908510" = "TOPICAL ANESTHETIC GASSES"
 "908599" = "TOPICAL ANESTHETIC COMBINATIONS"
 "908700" = "PIGMENTING-DEPIGMENTING AGENTS"
 "908710" = "PIGMENTING AGENTS"
 "908720" = "DEPIGMENTING AGENTS"
 "908799" = "DEPIGMENTING COMBINATIONS"
 "908800" = "AGENTS FOR FACIAL WRINKLES"
 "908860" = "AGENTS FOR FACIAL WRINKLES - RETINOIDS"
 "908900" = "GLABELLAR LINES (FROWN LINES) AGENTS"
 "909000" = "SCABICIDES & PEDICULOCIDES"
 "909010" = "NIT REMOVERS"
 "909099" = "SCABICIDE COMBINATIONS"
 "909200" = "SUNSCREENS"
 "909300" = "SCAR TREATMENT PRODUCTS"
 "909400" = "WOUND CARE PRODUCTS"
 "909420" = "WOUND CICATRIZING (SCARRING) AGENTS"
 "909430" = "WOUND CLEANSERS/DECUBITUS ULCER THERAPY"
 "909440" = "WOUND DRESSINGS"
 "909450" = "WOUND CARE - GROWTH FACTOR AGENTS"
 "909460" = "WOUND CARE - THROMBOLYTIC AGENTS"
 "909470" = "WOUND CARE SUPPLIES"
 "909480" = "SKIN TISSUE REPLACEMENTS"
 "909499" = "WOUND CARE COMBINATIONS"
 "909500" = "POISON IVY PRODUCTS"
 "909599" = "POISON IVY PRODUCT COMBINATIONS"

"909600" = "TOPICAL VASOPROTECTANTS"
"909699" = "TOPICAL VASOPROTECTANT COMBINATIONS"
"909700" = "MISC. TOPICAL"
"909710" = "ASTRINGENTS"
"909715" = "SKIN CLEANSERS"
"909720" = "SKIN PROTECTANTS"
"909730" = "SOAPS"
"909740" = "SHAMPOOS"
"909750" = "POWDERS"
"909760" = "SKIN OILS"
"909770" = "LUBRICANTS"
"909780" = "EYELID CLEANSERS"
"909799" = "MISC. TOPICAL COMBINATIONS"
"909800" = "PODIATRIC PRODUCTS"
"909899" = "PODIATRIC PRODUCT - COMBINATIONS"
"909900" = "MISC. DERMATOLOGICAL PRODUCTS"
"910000" = "R E S E R V E D"

GROUPS 92-99 MISCELLANEOUS PRODUCTS

"920000" = "ANTISEPTICS & DISINFECTANTS"
"921000" = "CHLORINE ANTISEPTICS"
"921099" = "CHLORINE ANTISEPTIC COMBINATIONS"
"922000" = "IODINE ANTISEPTICS"
"922099" = "IODINE ANTISEPTIC COMBINATIONS"
"923000" = "MERCURY ANTISEPTICS"
"924000" = "SILVER ANTISEPTICS"
"925000" = "WATER PURIFIERS"
"928000" = "DISINFECTANTS"
"929900" = "ANTISEPTIC COMBINATIONS"
"930000" = "ANTIDOTES"
"931000" = "CHELATING AGENTS"
"932000" = "BENZODIAZEPINE ANTAGONISTS"
"934000" = "NARCOTIC ANTAGONISTS"
"908000" = "TOPICAL ANTIDOTES"
"939900" = "ANTIDOTE KITS"
"940000" = "DIAGNOSTIC PRODUCTS"
"941000" = "DIAGNOSTIC REAGENTS"
"941010" = "INFECTION TESTS"
"941075" = "CONTROL REAGENTS"
"941098" = "DIAGNOSTIC TEST COMBINATIONS"
"941099" = "MULTIPLE URINE TESTS"
"942000" = "DIAGNOSTIC DRUGS"
"942099" = "DIAGNOSTIC DRUG COMBINATIONS"
"943000" = "DIAGNOSTIC BIOLOGICALS"
"943099" = "MULTIPLE SKIN TESTS"

"943500" = "DIAGNOSTIC RADIOPHARMACEUTICALS"
 "943540" = "DIAGNOSTIC RADIOPHARMACEUTICALS - ENDOCRINE"
 "943565" = "DIAGNOSTIC RADIOPHARMACEUTICALS - PROSTATIC"
 "944000" = "RADIOGRAPHIC CONTRAST MEDIA"
 "944010" = "RADIOGRAPHIC CONTRAST MEDIA - BARIUM"
 "944020" = "RADIOGRAPHIC CONTRAST MEDIA - IODINATED"
 "944030" = "RADIOGRAPHIC CONTRAST MEDIA - IRON OXIDE"
 "945000" = "NON-RADIOGRAPHIC CONTRAST MEDIA"
 "945099" = "NON-RADIOGRAPHIC CONTRAST MEDIA COMBINATIONS"
 "946000" = "DIAGNOSTIC PRODUCTS, MISC."
 "946060" = "DIAGNOSTIC SUPPLIES"
 "950000" = "ALTERNATIVE MEDICINES"
 "960000" = "CHEMICALS"
 "961000" = "ACIDS, BASES, & BUFFERS"
 "961010" = "ACIDS"
 "961020" = "BASES"
 "961030" = "BUFFERS"
 "962000" = "LIQUIDS"
 "962010" = "SOLVENTS"
 "962020" = "FIXED OILS"
 "962025" = "ESSENTIAL OILS"
 "962099" = "LIQUID COMBINATIONS"
 "963000" = "SOLIDS"
 "963010" = "ADDITIONAL SOLIDS"
 "963099" = "SOLID COMBINATIONS"
 "964000" = "SEMI-SOLIDS"
 "970000" = "MEDICAL DEVICES"
 "970500" = "PARENTERAL THERAPY SUPPLIES"
 "970510" = "NEEDLES & SYRINGES"
 "970520" = "IV SETS/TUBING"
 "970525" = "PARENTERAL CATHETERS"
 "970530" = "BLOOD ADMINISTRATION SETS"
 "970540" = "SUBCUTANEOUS ADMINISTRATION SUPPLIES"
 "970800" = "CARDIOLOGY SUPPLIES"
 "970820" = "CARDIAC CATHETERS"
 "970830" = "CARDIAC DEFIBRILLATORS"
 "970860" = "CARDIAC PACEMAKERS"
 "971000" = "RESPIRATORY THERAPY SUPPLIES"
 "971010" = "NEBULIZERS"
 "971015" = "PEAK FLOW METERS"
 "971020" = "HUMIDIFIERS"
 "971027" = "STEAM INHALERS"
 "971030" = "VAPORIZERS"
 "971050" = "TRACHEOSTOMY CARE & SUPPLIES"
 "971200" = "RESPIRATORY AIDS"
 "971210" = "MASKS"

"971240" = "AIR CLEANERS"
 "971500" = "GI-GU OSTOMY - IRRIGATION SUPPLIES"
 "971505" = "CATHETERS"
 "971510" = "OSTOMY SUPPLIES"
 "971520" = "INCONTINENCE SUPPLIES"
 "971525" = "IRRIGATION - TYPE SYRINGES"
 "971530" = "URINARY DRAINAGE & IRRIGATION SUPPLIES"
 "971600" = "HEMODIALYTICS & HEMOFILTRATES"
 "971610" = "HEMODIALYSIS"
 "971700" = "PERITONEAL DIALYSIS"
 "971710" = "PERITONEAL DIALYSIS - CATHETERS"
 "971720" = "PERITONEAL DIALYSIS - DEVICES"
 "972000" = "DIABETIC SUPPLIES"
 "972010" = "INSULIN ADMINISTRATION SUPPLIES"
 "972020" = "GLUCOSE MONITORING TEST SUPPLIES"
 "972023" = "GLUCOSE MONITORING & INSULIN ADMINISTRATION SUPPLIES"
 "972025" = "GLUCOSE/KETONE MONITORING TEST SUPPLIES"
 "972500" = "ENTERAL NUTRITION SUPPLIES"
 "972510" = "FEEDING TUBES"
 "973000" = "BANDAGES - DRESSINGS - TAPE"
 "973010" = "ADHESIVE BANDAGES"
 "973015" = "COMPRESSION BANDAGES"
 "973020" = "GAUZE BANDAGES"
 "973030" = "GAUZE PADS & DRESSINGS"
 "973040" = "ADHESIVE TAPE"
 "973050" = "NASAL DILATORS"
 "973070" = "SCAR TREATMENTS"
 "973400" = "FIXED (RIGID) BANDAGES/SUPPORTS & ACCESSORIES"
 "973440" = "CASTS, SPLINTS & ACCESSORIES"
 "973500" = "ELASTIC BANDAGES & SUPPORTS"
 "973600" = "SURGICAL SUPPLIES"
 "973650" = "SURGICAL SCALPELS & BLADES"
 "973660" = "SUTURES"
 "973670" = "SURGICAL ORTHOPEDIC DEVICES"
 "973700" = "HEATING/COOLING AIDS"
 "973710" = "HEATING PADS"
 "973720" = "HOT PACKS"
 "973750" = "COLD PACKS"
 "973790" = "HOT/COLD COMBINATION THERAPY AIDS"
 "973900" = "BACK PLASTERS"
 "974000" = "CONTRACEPTIVES"
 "974010" = "CONDOMS - MALES"
 "974015" = "CONDOMS - FEMALE"
 "974018" = "CERVICAL CAPS"
 "974020" = "DIAPHRAGMS"
 "974030" = "IUD'S"

"974040" = "CONTRACEPTIVE SPONGE"
 "974200" = "FERTILITY MONITORING TEST SUPPLIES"
 "974500" = "FEMALE PERSONAL CARE PRODUCTS"
 "974510" = "SANITARY NAPKINS & TAMPONS"
 "974520" = "DOUCHE SUPPLIES"
 "974700" = "IMPOTENCE AIDS"
 "974750" = "IMPOTENCE AIDS - MALE"
 "974760" = "SEXUAL DYSFUNCTION DEVICES - FEMALE"
 "975000" = "ORAL HYGIENE PRODUCTS"
 "975005" = "DENTAL SUPPLIES"
 "975010" = "DENTURE CARE PRODUCTS"
 "975020" = "TOOTHBRUSHES - FLOSS"
 "975030" = "DENTIFRICES"
 "975040" = "DENTAL WHITENING PRODUCTS"
 "975500" = "INFANT CARE PRODUCTS"
 "975510" = "FEEDING SUPPLIES"
 "975520" = "DIAPERS"
 "975530" = "NURSING PADS"
 "975700" = "AUDITORY SUPPLIES"
 "975740" = "HEARING AID SUPPLIES-BATTERIES"
 "976000" = "OPTICAL SUPPLIES"
 "976010" = "CONTACT LENS CARE SUPPLIES"
 "976020" = "EYEGLOSS CARE SUPPLIES"
 "976030" = "EYE PATCHES"
 "976500" = "DURABLE MEDICAL EQUIPMENT"
 "977000" = "MISC. DEVICES"
 "977010" = "THERMOMETERS"
 "977020" = "DISPOSABLE GLOVES"
 "977030" = "APPLICATORS, COTTON BALLS, ETC."
 "977040" = "RUBBER GOODS"
 "977050" = "NERVE STIMULATORS"
 "977053" = "BLOOD COAGULATION TEST SUPPLIES"
 "977055" = "NEUROLOGICAL DIAGNOSTIC SUPPLIES"
 "977060" = "OPHTHALMIC PRESSURE MONITORS"
 "977070" = "RAZORS AND BLADES"
 "977080" = "SPONGES"
 "977085" = "DRUG APPLICATION PRODUCTS"
 "977090" = "ORAL DOSING DEVICES"
 "977095" = "ACUPRESSURE THERAPY SUPPLIES"
 "977500" = "BLOOD PRESSURE DEVICES"
 "977599" = "BLOOD PRESSURE DEVICE COMBINATIONS"
 "978000" = "FOOT CARE PRODUCTS"
 "978099" = "FOOT CARE PRODUCT - COMBINATIONS"
 "978500" = "FIRST AID KITS"
 "980000" = "PHARMACEUTICAL ADJUVANTS"
 "980500" = "ALKALIZING AGENTS"

"981000" = "ANITMICROBIAL AGENTS"
 "982000" = "ANTIOXIDANTS"
 "982500" = "PRESERVATIVES OTHER"
 "983000" = "COLORING AGENTS"
 "983300" = "FLAVORING AGENTS"
 "983500" = "PHARMACEUTICAL EXCIPIENTS"
 "983510" = "EXTERNAL VEHICLE INGREDIENTS"
 "983600" = "INTERNAL VEHICLE INGREDIENTS/AGENTS"
 "983650" = "THICKENING AGENTS"
 "983700" = "SURFACTANTS"
 "984000" = "LIQUID VEHICLE"
 "984010" = "PARENTERAL VEHICLES"
 "984020" = "ORAL VEHICLES"
 "984030" = "EXTERNAL VEHICLES"
 "984090" = "MISC. VEHICLES"
 "986000" = "SEMISOLID VEHICLE"
 "986500" = "DELIVERY DEVICES"
 "986520" = "TRANSDERMAL PATCHES"
 "987000" = "GELATIN CAPSULES (EMPTY)"
 "987100" = "NON GELATIN CAPSULES (EMPTY)"
 "988000" = "PLACEBOS"
 "989000" = "PHARMACEUTICAL ADJUVANTS MISCELLANOUS"
 "989030" = "BUFFER REFERENCE STANDARDS"
 "990000" = "ASSORTED CLASSES"
 "991000" = "CARDIOPLEGIC SOLUTION"
 "992000" = "CHELATING AGENTS"
 "993000" = "COLLAGEN IMPLANT"
 "993200" = "CYTOPROTECTIVE AGENTS"
 "993250" = "LAZARIDS"
 "993500" = "ENZYMES"
 "993900" = "IMMUNOMODULATORS"
 "993920" = "ANTILEPTICS"
 "994000" = "IMMUNOSUPPRESSIVE AGENTS"
 "994020" = "CYCLOSPORINE ANALOGS"
 "994025" = "IMMUNE GLOBULIN IMMUNOSUPPRESSANTS"
 "994030" = "INOSINE MONOPHOSPHATE DEHYDROGENASE INHIBITORS"
 "994040" = "MACROLIDE IMMUNOSUPPRESSANTS"
 "994050" = "MONOCLONAL ANTIBODIES"
 "994060" = "PURINE ANALOGS"
 "994070" = "PYRIMIDINE ANALOGS"
 "994500" = "K REMOVING RESIN"
 "994700" = "LYMPHATIC AGENTS"
 "994750" = "LYMPHEDEMA AGENTS"
 "995000" = "PROSTAGLANDINS"
 "996500" = "SCLEROSING AGENTS"
 "996599" = "SCLEROSING COMBINATIONS"

"997000" = "PERITONEAL DIALYSIS SOLUTIONS"
"997500" = "IRRIGATION SOLUTIONS"
"998000" = "ORGAN PRESERVATION SOLUTION"
"998500" = "MISC. NATURAL PRODUCTS"
"998700" = "HOMEOPATHIC PRODUCTS"
"999000" = "NOT CLASSIFIED"
"999030" = "UNCLASSIFIED OTC PRODUCT"